

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: September 2, 2004, 18:28:29 ; Search time 2829 Seconds
(without alignments)
3998.774 Million cell updates/sec

Title: US-10-063-731-118
Perfect score: 1357
Sequence: 1 MSTTCQVAFLLSILGLAG.....DGGARTEDVQSPKSHDYV 261

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10063731/runat_01092004_155040_18714/app_query.fasta_1.455
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -RND=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10063731_@CGN_1_1_2521_@runat_01092004_155040_18714 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_hgt_hum:*
31: em_hgt_inv:*
32: em_hgt_other:*
33: em_hgt_mus:*
34: em_hgt_pln:*
35: em_hgt_rod:*
36: em_hgt_mam:*
37: em_hgt_vrt:*
38: em_sy:*
39: em_hgtgo_hum:*
40: em_hgtgo_mus:*
41: em_hgtgo_other:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1357	100.0	816	9	AF221069	AF221069 Homo sapi
2	1357	100.0	982	6	BD237992	BD237992 Gastric p
3	1357	100.0	1869	6	AX472954	AX472954 Sequence
4	1357	100.0	2121	6	AX092386	AX092386 Sequence
5	1357	100.0	2121	6	AX376288	AX376288 Sequence
6	1357	100.0	2121	6	AX697257	AX697257 Sequence
7	1357	100.0	2121	9	AY358479	AY358479 Homo sapi
8	1264	93.1	796	9	AF349452	AF349452 Homo sapi
9	1264	93.1	3344	9	AK098474	AK098474 Homo sapi
10	1235.5	91.0	2786	10	AF221068	AF221068 Mus muscu
11	1217.5	89.7	1409	10	AF349450	AF349450 Mus muscu
12	1165.5	85.9	849	10	AF349451	AF349451 Mus muscu
13	1147.5	84.6	861	10	AF349453	AF349453 Mus muscu
14	1120	82.5	783	6	BD237993	BD237993 Gastric p
15	1107.5	81.5	2040	9	BC063845	BC063845 Homo sapi
16	1011.5	74.5	2388	5	BC060425	BC060425 Xenopus l
17	37	29.3	164953	9	AC016252	AC016252 Homo sapi
18	394	29.0	423	6	AX380609	AX380609 Sequence
19	394	29.0	423	6	AX380725	AX380725 Sequence
20	389	28.7	145477	2	AC010810	AC010810 Homo sapi
21	388	28.6	636	6	E31608	E31608 Tight junct
22	388	28.6	636	10	AF087825	AF087825 Mus muscu
23	388	28.6	894	10	BC008104	BC008104 Mus muscu
24	388	28.6	1235	10	BC050007	BC050007 Mus muscu
25	388	28.6	158198	2	AC012013	AC012013 Homo sapi
26	385	28.4	888	10	AF486651	AF486651 Mus muscu
27	382	28.2	2859	6	AX833932	AX833932 Sequence
28	382	28.2	2859	6	AX096063	AX096063 Homo sapi
29	375.5	27.7	690	6	AX921675	AX921675 Sequence
30	375.5	27.7	1616	9	BC030524	BC030524 Homo sapi
31	373	27.5	245225	2	AC098329	AC098329 Rattus no
32	373	27.5	247885	2	AC097923	AC097923 Rattus no
33	370	27.3	208814	2	AC134606	AC134606 Mus muscu
34	370	27.3	261333	2	AC138280	AC138280 Mus muscu
35	368.5	27.2	1207	9	HSA011497	HSA011497 Homo sapi
36	368.5	27.2	1253	6	AR400627	AR400627 Sequence
37	368.5	27.2	1253	6	AX013131	AX013131 Sequence
38	368.5	27.2	1253	6	BD206312	BD206312 Human nuc
39	368.5	27.2	1546	6	BD234896	BD234896 Human apo
40	368.5	27.2	1711	6	BD063242	BD063242 Secreted
41	368.5	27.2	1787	6	AX098214	AX098214 Sequence
42	368	27.1	681	6	AX276968	AX276968 Sequence
43	367	27.0	580	10	AF249889	AF249889 Mus muscu
44	364.5	26.9	678	6	AX083410	AX083410 Sequence
45	364.5	26.9	1095	6	AX083420	AX083420 Sequence

ALIGNMENTS

RESULT 1

```
AF221069      816 bp      mRNA      linear      PRI 10-OCT-2001
LOCUS      Homo sapiens Claudin-18 mRNA, complete cds.
DEFINITION      AF221069
ACCESSION      AF221069
VERSION      AF221069.1 GI:6715517
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 816)
Niimi,T., Nagashima,K., Ward,J.M., Minoo,P., Zimonjic,D.B.,
Popescu,N.C. and Kimura,S.
claudin-18, a novel downstream target gene for the T/EBP/NKX2.1
homeodomain transcription factor, encodes lung- and
stomach-specific isoforms through alternative splicing
Mol. Cell. Biol. 21 (21), 7380-7390 (2001)
21470315
JOURNAL
MEDLINE
PubMed
11585919
REFERENCE      2 (bases 1 to 816)
AUTHORS      Niimi,T. and Kimura,S.
TITLE      Direct Submision
JOURNAL      Submitted (04-JAN-2000) Laboratory of Metabolism, DBS, National
Cancer Institute, NIH, 9000 Rockville Pike, Bethesda, MD 20892, USA
FEATURES
source
1..816
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
10..795
/codon_start=1
/product="Claudin-18"
/protein_id="AAF26448.1"
/db_xref="GI:6715518"
translation="MSTTTCQVAFLLSILGLACIAATGMDMSTODLYDNPVTSVP
QYGLMPSVQSSGFTCEPFIYILGLPAMLQAVRALMIVGIVLGAIGLIVSLPALK
CIRGSMDSKANNILTSGLIMFIVSGLIAGVSVFANMLVTFNFMSTANNMTGMGG
MVQTVQRTYFGAALFVGWVAGGLTILGVMNCIACRGLAPEETNYKAVSHAGSHV
AYKPGGFKASTGFGSNTKNKKIYDGGARTDEVOYSPSKHDYV"
ORIGIN
Alignment Scores:
Pred. No.:      2.75e-126      Length:      816
Score:      1357.00      Matches:      261
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      100.00%      Indels:      0
DB:      9      Gaps:      0
US-10-063-731-118 (1-261) x AF221069 (1-816)
Qy      1 MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db      10 ATGTCCACCACCATGCGAAGTGGTGGCGTTCCTCTGTCATCTCGGGCTGGCGGC 69
Qy      21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyAspAsnProVal 40
Db      70 TGCATCGCGGCACCGGATGGACATGTGGAGCACCAGGACCTGTACGACCAACCCGTC 129
Qy      41 ThrSerValPheGlnTyGluGlyLeuTrpArgSerCysValAlaGlnSerSerGlyPhe 60
Db      130 ACCTCCGTGTTCCAGTACGAGGCGCTTGGAGGAGCTGCGTGAGCGAGATTACGGTTC 189
Qy      61 ThrGluCysArgProTyTrpPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db      190 ACCGAATGCAGGCCCTATTTCCACCATCTCGGACTTCCAGCCATGCTCGAGGACGTGCGA 249
Qy      81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuValSerIlePheAla 100
Db      250 GCCCTGATGATCGTAGGCATGCTCTGGGTGCCATTGGCTCTCTGGTATCCATCTTTGCC 309
Qy      101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
```

```
Db      310 CTGAATGATCATCCGCAITGGCAGCATGGAGGACTCTGCCAAAGCCAAACATGACACTGACC 369
Qy      121 SerGlyIleMetPheIleValSerCysLeuCysAlaIleAlaGlyValSerValPheAla 140
Db      370 TCCGGGATCATGTTCAITGTCTCAGGTCTTTGTGCAATTGCTGGAGTGTCTGTGTGGCTGCC 429
Qy      141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyThrGlyMetGlyGly 160
Db      430 AACATGCTGGTGACTTAACCTTGGATGTCACAGATACATGATACACCGGCATGGGTGGG 489
Qy      161 MetValGlnThrValGlnThrArgTyThrPheGlyAlaAlaLeuPheValGlyTrpVal 180
Db      490 ATGGTGCAGACTGTTTCAGACAGGTACACATTTGGTGGCGGTCTGTCTGGGTGGGTGC 549
Qy      181 AlaGlyGlyLeuThrLeuIleGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db      550 GCTGGAGGCTCTCACACTAAATTGGGGGTGTGATGATGTGCTGCTGCGGGCGCTGGCA 609
Qy      201 ProGluGlnThrAsnTyLysAlaValSerTyHisAlaSerGlyHisSerValAlaTy 220
Db      610 CCAGAAGAAACCACTACAAAGCCGTTCTTATCATGCTCAGGCCACAGTGTTCCTTAC 569
Qy      221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysLysIle 240
Db      670 AAGCTGGAGGCTTCAAGGCCAGCAGCTGGCTTGGGTCCAAACCAAAAACAAGAAGATA 729
Qy      241 TyrAspGlyGlyAlaAlaArgThrGluAspGluValGlnSerTyProSerLysHisAspTy 260
Db      730 TACGATGGAGGTGCCGCCACAGAGCAGGATGATCTTATCTTATCTTCCAGCAGCTAT 789
Qy      261 Val 261
Db      790 GTG 792
RESULT 2
BD237992
LOCUS      Gastric polypeptide ZSIG28.
DEFINITION      BD237992
ACCESSION      BD237992.1 GI:33047762
VERSION      JP 2002524103-A/1.
KEYWORDS      Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 982)
Sheppard,P.O. and Foley,K.P.
Gastric polypeptide ZSIG28
Patent: JP 2002524103-A 1 06-AUG-2002;
ZYMOGENETICS INC
COMMENT      OS Homo sapiens (human)
PN JP 2002524103-A/1
PD 06-AUG-2002
PF 14-SEP-1999 JP 2000570197
PR 16-SEP-1998 US 09/154444
PI PAUL O SHEPPARD, KEVIN P FOLEY
PC C12N15/09,A61K38/00,A61K39/395,A61K39/395,A61P1/04,A61P1/14,
PC A61P3/08,
PC A61P5/50,A61P31/04,A61P31/10,A61P35/00,C07K14/47,C07K16/18, PC
C12N1/15,
PC C12N1/19,C12N1/21,C12N5/10,C12P21/02,C12P21/08,C12Q1/02 PC
C12Q1/68,G01N33/15,
PC G01N33/50,G01N33/53,G01N33/577,C12N15/00,C12N5/00,A61K37/02 CC
Gastric polypeptide ZSIG28
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..982
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
```

Alignment Scores:		3.49e-126	Length:	982
Pred. No.:	1	1357.00	Matches:	261
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	100.00%	Indels:	0	
DB:	6	Gaps:	0	
US-10-063-731-118 (1-261) x BD237992 (1-982)				
Qy	1	MetSerThrThrCysGlnValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly	20	
Db	70	ATGTCACACACACATGCAAGTGTGGGTTCTCTCTGTCATCTCTGGGGCTGGCGGC	129	
Qy	21	CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyAspAsnProVal	40	
Db	130	TGCATCGCGGCACCGGATGGACATGTGGAGCACCAGGACCTGTACGACAAACCCGTC	189	
Qy	41	ThrSerValPheGlnTyGluGlyLeuTrpArgSerCysValArgGlnSerSerGlyPhe	60	
Db	190	ACCTCCGTTTCCAGTACGAAGGGCTCTGGAGAGCTGGGTGAGGACAGTTTACGGCTTC	249	
Qy	61	ThrGluCysArgProTyPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg	80	
Db	250	ACCGAATGAGGCCCTATTTCACCATCTCTGGGACTTCAGGCCATGCTGAGGACGTGCGA	309	
Qy	81	AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla	100	
Db	310	GCCCTGATGATCGTAGGCATCTCTGGGTGGCAATTTGGCTCTCTGATTCATCTTTGCC	369	
Qy	101	LeuLysCysIleAArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr	120	
Db	370	CTGAATGATGATCCGATTTGGCAGTGGAGGACTCTGCCAAGGCCACATGACATGACC	429	
Qy	121	SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla	140	
Db	430	TCCGGATCATCTTATTGTCAGGTCTTTGTGCAATTTGCTGGAGTGTCTGTGTTGCC	489	
Qy	141	AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyThrGlyMetGlyGly	160	
Db	490	AACATGCTGGTGACTTAACCTTCTGGATGTCCACAGCTAACATGTACACCGGCGATGGG	549	
Qy	161	MetValGlnThrValGlnThrArgTyThrPheGlyAlaAlaLeuPheValGlyTrpVal	180	
Db	550	ATGGTGACAGTGTTCAGACAGGATACATATTGGGCGGCTCTGTGTGGGTGGGTC	609	
Qy	181	AlaGlyGlyLeuThrLeuIleGlyValMetMetCysIleAlaCysArgGlyLeuAla	200	
Db	610	GCTGGAGGCTCACACTAATTGGGGGTGTGATGTGATGTCATCGCTTCCGGGGCTGGCA	669	
Qy	201	ProGluGluThrAsnTyrlsAlaValSerTyrlsAlaSerGlyHisSerValAlaTy	220	
Db	670	CCAGAAGAACAACCACTACAAAGCCGTTCTTATCATGCTCAGGCCACAGTGTGCTTAC	729	
Qy	221	LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysIle	240	
Db	730	AAGCTTGGAGGCTTCAAGCCAGCATCTGGCTTTGGGTCCACACCAAAACCAAGAGATA	789	
Qy	241	TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrlsProSerLysHisAspTy	260	
Db	790	TACGATGGAGGTGCCCGCACAGAGGACGAGGTACATCTTATCTTCCAGGACGACTAT	849	
Qy	261	Val 261		
Db	850	GTG 852		
RESULT 3				
AX472954				
LOCUS				
DEFINITION				
Sequence 3 from Patent WO0218576.				
ACCESSION				
AX472954				
VERSION				
AX472954.1 GI:22207741				
KEYWORDS				
linear				
PAT 09-AUG-2002				

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1
TITLE	Chen, S. Y., Macina, R. A., Sun, Y. and Recipon, H.
JOURNAL	Compositions and methods relating to lung specific genes
FEATURES	Patent: WO 0218576-A 3 07-MAR-2002;
Location/Qualifiers	Diadexis, Inc. (US)
source	1. .1869
/organism="Homo sapiens"	/mol_type="unassigned DNA"
/db_xref="taxon:9606"	
ORIGIN	
Alignment Scores:	7.96e-126 Length: 1869
Pred. No.:	1357.00 Matches: 261
Score:	100.00% Conservative: 0
Percent Similarity:	100.00% Mismatches: 0
Best Local Similarity:	100.00% Indels: 0
Query Match:	100.00% Gaps: 0
DB:	6
US-10-063-731-118 (1-261) x AX472954 (1-1869)	
Qy	1 MetSerThrThrCysGlnValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly
Db	83 ATGTCACACACACATGCAAGTGTGGGTTCTCTCTGTCATCTCTGGGGCTGGCGGC
Qy	21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyAspAsnProVal
Db	143 TGCATCGCGGCACCGGATGGACATGTGGAGCACCAGGACCTGTACGACAAACCCGTC
Qy	41 ThrSerValPheGlnTyGluGlyLeuTrpArgSerCysValArgGlnSerSerGlyPhe
Db	203 ACCTCCGTTTCCAGTACGAAGGGCTCTGGAGAGCTGTGGAGGACAGTTTACGGCTTC
Qy	61 ThrGluCysArgProTyPheThrIleLeuGlyLeuProAlaMetIleGlnAlaValArg
Db	263 ACCGAATGAGGCCCTATTTCACCATCTCTGGGACTTCAGGCCATGCTGCAGGACGTGCGA
Qy	81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla
Db	323 GCCCTGATGATCGTAGGCATCTCTGGGTGGCAATTTGGCTCTCTGATTCATCTTTGCC
Qy	101 LeuLysCysIleAArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr
Db	383 CTGAATGATGATCGGATTTGGCAGCATGGAGGACTCTGCCAAAGCCCAACATGACCTGACC
Qy	121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla
Db	443 TCCGGATCATGTTTATGTCAGGTCTTGTGCAATTTGGGTGGAGTGTCTGTGTTGCC
Qy	141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyThrGlyMetGlyGly
Db	503 AACATGCTGGTGACTTAACCTTCTGGATGTCCACAGCTAACATGTACACCGGATGGTGGG
Qy	161 MetValGlnThrValGlnThrArgTyThrPheGlyAlaAlaLeuPheValGlyTrpVal
Db	563 ATGGTGACAGTGTTCAGACCCAGGTACACATTTGGTGGGCTCTGTTCGTGGGCTGGGTC
Qy	181 AlaGlyGlyLeuThrLeuIleGlyValMetMetCysIleAlaCysArgGlyLeuAla
Db	623 GCTGGAGGCTCACACTAATTGGGGGTGTGATGTGTCATTCCTTCCGGGGCTGGCA
Qy	201 ProGluGluThrAsnTyrlsAlaValSerTyrlsAlaSerGlyHisSerValAlaTy
Db	683 CCAGAAGAACAACCACTACAAAGCCGTTCTTATCATGCTCAGGCCACAGTGTGCTTAC
Qy	221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysIle
Db	743 AAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCCAACCAAAACCAAGAGATA

```

Qy 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db 803 TACGATGGAGGTGGCCGACAGAGGACGAGGTACAACTCTTATCTTCCAAAGCAGACTAT 862

Qy 261 Val 261
Db 863 GTG 865

RESULT 4
AX092386 2121 bp DNA linear PAT 21-MAR-2001
LOCUS
DEFINITION Sequence 117 from Patent WO0116318.
ACCESSION AX092386
VERSION AX092386.1 GI:13444507
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Eaton,D.L., Filvaroff,E., Gerritsen,M.E., Goddard,A.,
Godowski,P.J., Grimaldi,C.J., Gurney,A.L., Watanabe,C.K. and
Wood,W.I.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0116318-A 117 08-MAR-2001;
Genentech, Inc. (US)
FEATURES
source
Location/Qualifiers
1. .2121
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 9,37e-126 Length: 2121
Score: 1357.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-063-731-118 (1-261) x AX092386 (1-2121)

Qy 1 MetSerThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 90 ATGTCCACCAACCATGCGAGTGGCGTTCCTCTGTCATCTCTGCGGCTGGCCGGC 149

Qy 21 CysIleAlaThrGlyMetAspMetTyrSerThrGlnAspLeuTyrAspAsnProVal 40
Db 150 TGCATCGCGGCCACCGGGATGGACATGTGGAGCACCAGGACCTGTACGACAAACCCGTC 209

Qy 41 ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe 60
Db 210 ACCTCCGTGTCCAGTACGAGGCGCTCTGGAGAGCTCGGTGAGCAGAGTTTCAGGTTTC 269

Qy 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 270 ACCGAATGCAGGCCCTATTTCCACATCTCGGACTTCAGCCATGTCTGAGGACGTGCGA 329

Qy 81 AlaLeuMetIleValGlyIleValIleValIleGlyLeuValSerIlePheAla 100
Db 330 GCCCTGATGATCGTAGGATCGTCTCGGTGGTCCATTCCTCTGATCCATCTTTGCC 389

Qy 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 390 CTGAATGTCATCCGATTCGAGCATGGAGGACTCTGCCAAGCAACATGACACTGAC 449

Qy 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 450 TCGGGGATCATGTTTCATTGTCTCAGGTCTTTGTGCAATTGTGAGTGTCTGTGTTTGGC 509

```

```

Qy 141 AsnMetLeuValThrAsnPheTyrMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db 510 AACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGCATGGTGGG 569

Qy 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTyrVal 180
Db 570 ATGGTGCAGACTGTTTCAGACAGGTACACATTTGGTGGCGCTCTGTTCGTGGGCTGGGTC 629

Qy 181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db 630 GCTGGAGGCTTCACACTAATTTGGGGGTGTGATGATGTGCATCGCTTCGGGGCTGGCA 689

Qy 201 ProGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
Db 690 CCAGAAGAAACCAACTACAAAGCCGTTCTTATCATGCTCAGGCCACAGTGTTCCTTAC 749

Qy 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysIle 240
Db 750 AAGCTTGAGGCTTCAAGGCCAGCAGCTGGCTTTGGGTCCACACCAAAAAACAAGAGATA 809

Qy 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db 810 TACATGGAGGTGCGCGACAGAGCAGAGGTACATCTTATCTTCCAGCACAGCTAT 869

Qy 261 Val 261
Db 870 GTG 872

RESULT 5
AX376288 2121 bp DNA linear PAT 01-MAR-2002
LOCUS
DEFINITION Sequence 355 from Patent WO0168848.
ACCESSION AX376288
VERSION AX376288.1 GI:19170535
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Baker,K.P., Chen,J., Desnoyers,L., Goddard,A., Godowski,P.J.,
Gurney,A.L., Pan,J., Smith,V., Watanabe,C.K., Wood,W.I. and
Zhang,Z.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0168848-A 355 20-SEP-2001;
Genentech, Inc. (US)
FEATURES
source
Location/Qualifiers
1. .2121
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 9,37e-126 Length: 2121
Score: 1357.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-063-731-118 (1-261) x AX376288 (1-2121)

Qy 1 MetSerThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 90 ATGTCCACCAACCATGCGAGTGGCGTTCCTCTGTCATCTCTGCGGCTGGCCGGC 149

Qy 21 CysIleAlaThrGlyMetAspMetTyrSerThrGlnAspLeuTyrAspAsnProVal 40
Db 150 TGCATCGCGGCCACCGGGATGGACATGTGGAGCACCAGGACCTGTACGACAAACCCGTC 209

Qy 41 ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe 60

```



```

Db      210  ACCTCGGTGTTCCAGTACGAGGGCTCTGGAGAGCTGCTGAGCAGAGTTCAGGCTTC 269
QY      61  ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db      270  ACCGAATGAGGCGCCCTATTTCACCATCTCTGGGACTTCCAGCCATGCTGCAGGACGTGCGA 329
QY      81  AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
Db      330  GGCCTGATGATCGTAGGCATCGCTCGGTGGTGCATTTGGCTCTGCTGATTCCTATCTTTGCC 389
QY      101  LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db      390  CTGAATGCATCGCATTCGACATGAGGAGTCTGCCAAGCCAAACATGACACTGACC 449
QY      121  SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db      450  TCCGGGATCATGTTTCATCTCTCAGGTCTTTGTGCAATTGTGCGAGTGTCTGTGTTGCC 509
QY      141  AsnMetLeuValThrAsnMetPheThrMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db      510  AACATGCTGCTGACTTAACCTCTGGATGTCACAGTAACTGATGTCATCGCTGCGGCTCTGCT 569
QY      161  MetValGlnThrValGlnThrArgTyrThrPheGlyAlaLeuPheValGlyTyrPVal 180
Db      570  ATGGTGCAGCTGTTCCAGACAGGTACACATTTGGTGGGCTCTGTTCTGCGGCTGGGTC 629
QY      181  AlaGlyGlyLeuThrIleGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db      630  GCTGGAGGCTCACACTAAATGCGGGTGTGATGATGTCATCGCTGCGGCGGCGGCTGGCA 689
QY      201  ProGluGlnThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
Db      690  CCAGAGAAGAACCAACTACAAAGCCGTTCTTATCATGCTCTCAGCCACAGTGTGTGCTTAC 749
QY      221  LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysIle 240
Db      750  AGCCCTGGAGGCTTCAAGCCAGCAGCTGCTTTGGTCCCAACACCAACCAAGAGATA 809
QY      241  TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db      810  TACGATGGAGGTGCCCGCACAGAGGACGAGGTACAACTCTTATCCTTCCCAAGCAGCACTAT 869
QY      261  Val 261
Db      870  GTG 872

```

```

RESULT 6
LOCUS   AX697257
DEFINITION Sequence 325 from Patent WO0078961.
ACCESSION AX697257
VERSION   AX697257.1 GI:29498417
KEYWORDS
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1  Ferrara,N., Stewart,T.A., Williams,P.M., Baker,K.P., Desnoyers,L.,
      Eaton,D.L., Gao,W.Q., Pan,J., Botstein,D., Fong,S., Goddard,A.,
      Godowski,P.J., Gurney,A.L., Smith,V., Tumas,D., Wood,W.I.,
      Grimaldi,C.J., Hillan,K.J., Paoni,N.F., Roy,M.A., and Watanabe,C.K.
      Secreted and transmembrane polypeptides and nucleic acid encoding
      the same
      Patent: WO 0078961-A 325 28-DEC-2000;
      Genentech Inc. (US)
FEATURES
  source             1..2121
                    /organism="Homo sapiens"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:9606"
ORIGIN

```

```

Alignment Scores:
Pred. No.:      9,37e-126      Length:      2121
Score:          1357.00      Matches:    261
Percent Similarity: 100.00%      Conservative: 0
Best local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:    0
DB:             6           Gaps:      0
US-10-063-731-118 (1-261) x AX697257 (1-2121)

```

```

QY      1  MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db      90  ATGTCACCAACCAATGCAAGTGTGGCTTCTCTCTGCTCCATCTGTCGCGCTGGCGGC 149
QY      21  CysIleAlaAlaThrGlyMetAspMetTyrSerThrGlnAspLeuTyrAspAsnProVal 40
Db      150  TGCATCGGGGCGCACCGGGATGGACATGGAGCACCCAGGACCTGTACGACAAACCCCGTC 209
QY      41  ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe 60
Db      210  ACCTCGGTGTTCCAGTACGAAGGGCTCTGGAGAGCTGGTCGAGGACAGTTCAGGCTTC 269
QY      61  ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db      270  ACCGAATGAGGCGCCCTATTTCACCATCTCTGGGACTTCCAGCCATGCTGCAGGACGTGCGA 329
QY      81  AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
Db      330  GGCCTGATGATCGTAGGCATCGCTCGGTGGTGCATTTGGCTCTCTGATTCCTATCTTTGCC 389
QY      101  LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db      390  CTGAATGCATCGCATTCGACATGAGGAGTCTGCCAAGCCAAACATGACACTGACC 449
QY      121  SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db      450  TCCGGGATCATGTTTCATCTCTCAGGTCTTTGTGCAATTGTGCGAGTGTCTGTGTTGCC 509
QY      141  AsnMetLeuValThrAsnMetPheThrMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db      510  AACATGCTGCTGACTTAACCTCTGGATGTCACAGTAACTGATGTCATCGCTGCGGCGGCTGGCA 569
QY      161  MetValGlnThrValGlnThrArgTyrThrPheGlyAlaLeuPheValGlyTyrPVal 180
Db      570  ATGGTGCAGCTGTTCCAGACAGGTACACATTTGGTGGGCTCTGTTCTGCGGCTGGGTC 629
QY      181  AlaGlyGlyLeuThrIleGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db      630  GCTGGAGGCTCACACTAAATGCGGGTGTGATGATGTCATCGCTGCGGCGGCGGCTGGCA 689
QY      201  ProGluGlnThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
Db      690  CCAGAGAAGAACCAACTACAAAGCCGTTCTTATCATGCTCTCAGCCACAGTGTGTGCTTAC 749
QY      221  LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysIle 240
Db      750  AGCCCTGGAGGCTTCAAGCCAGCAGCTGCTTTGGTCCCAACACCAACCAAGAGATA 809
QY      241  TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db      810  TACGATGGAGGTGCCCGCACAGAGGACGAGGTACAACTCTTATCCTTCCCAAGCAGCACTAT 869
QY      261  Val 261
Db      870  GTG 872

```

```

RESULT 7
LOCUS   AY358479
DEFINITION Homo sapiens clone DNA73734 CLDN18 (UNQ778) mRNA, complete cds.
ACCESSION AY358479
VERSION  AY358479.1 GI:37182080

```

```

KEYWORDS  FLI_CDNA.
SOURCE      Homo sapiens (human)
ORGANISM    Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 2121)
AUTHORS     Clark,H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Brush,J.,
            Chen,J., Chow,B., Chui,C., Crowley,C., Curreli,B., Deuel,B.,
            Dowd,P., Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Hass,P.E.,
            Heidens,S., Huang,A., Kim,H.S., Klimowski,L., Jin,Y., Johnson,S.,
            Lee,J., Lewis,L., Liao,D., Mark,M., Robbie,E., Sanchez,C.,
            Schoenfeld,J., Seshagiri,S., Simmons,L., Singh,J., Smith,V.,
            Stinson,J., Vagts,A., Vandlen,R., Watanabe,C., Wicand,D., Woods,K.,
            Xie,M.H., Yansura,D., Yi,S., Yu,G., Yuan,J., Zhang,M., Zhang,Z.,
            Goddard,A., Wood,W.I. and Godowski,P.
TITLE       The Secreted Protein Discovery Initiative (SPDI), a Large-Scale
            Effort to Identify Novel Human Secreted and Transmembrane Proteins:
            A Bioinformatics Assessment
JOURNAL     Genome Res. 13 (10), 2265-2270 (2003)
PUBMED      12975309
REFERENCE   2 (bases 1 to 2121)
AUTHORS     Clark,H.F.
TITLE       Direct Submission
JOURNAL     Submitted (01-AUG-2003) Department of Bioinformatics, Genentech,
            Inc., 1 DNA Way, South San Francisco, CA 94080, USA
FEATURES    source
            location/Qualifiers
            1..2121
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="DNA73734"
            1..2121
            /locus_tag="UNQ778"
            /locus_tag="UNQ778"
            /note="P01572"
            /codon_start=1
            /product="CLDN18"
            /protein_id="AAQ88843.1"
            /db_xref="GI:37182081"
            /translation="MSTTCQVAFLLSILGLACIAATGMMDMSTODLYDNPVSVF
            QYGLRSCVROSGFTCEPCPYFTILGLPAMLOAVRLMIVGLGAIGLLVSTFALK
            CIRGSMEDSAKANMTLTSGIMFIVSLGLACAGVSFANMLVTNFWNSTANVTGMGG
            MVQTVQTRYTFGAALFVGVAGGLTLGGVMCTACGLAPEETNYKAVSYHAGSHV
            AYRPGGFKASTGSGNTKNNKIYDGGARTEDVQSYPSKHDYV"

gene
CDS
            81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
            330 GCCCTGATGATCGTAGGCATCGTCTGGGTGCCATTCCTGCTGCTCCATCTTTGCC 389
            101 LeuLysCysIleAeGlyIleGlySerMetClnuSpSerAlaLysAlaAsnMetThrLeuThr 120
            390 CTGAATGTCATCCCATTTGGCAGCATGGAGACACTCTGCCAAGCCAAACATGACACTGACC 449
            121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
            450 TCCGGGATCATGTTTCATGTCTCAGGCTTTGTGCAATTTGCTGGAGTGTCTGTGTTGCC 509
            141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
            510 AACATGCTGGTGACTTAACCTCTGGATGTCCACAGCTAACATGATACACCGCATGGGTGG 569
            161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal 180
            570 ATGGTGCAGACTGTTTCAGACAGGATACATTTGGTGGGCTCTGTTCGTGGGTGGGTGC 629
            181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetCysIleAlaCysArgGlyLeuAla 200
            630 GCTGGAGGCGCTCACACTAAATTTGGGGTGTGATGATGTGATCGCTGCGGGGCTGGCA 689
            201 ProGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
            690 CCAGAAGAAACCAACTACAAAGCGTTTCTTATCATGCTCCTCAGGCCACAGTGTTCCTAC 749
            221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysIle 240
            750 AAGCTCTGGAGCTTCAAGGCCAGCACCTGGCTTTGGGTCACACCAACCAACCAAGAGATA 809
            241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
            810 TACGATGGAGGTGCGCCGACAGAGAGAGAGTACATCTTATCTTCCACGACGACTAT 869
            261 Val 261
            870 GTG 872

RESULT 8
AF349452 Homo sapiens claudin-18A2.1 mRNA, complete cds, alternatively
LOCUS     spliced.
DEFINITION AF349452
ACCESSION AF349452.1 GI:16224168
VERSION    AF349452
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 796)
AUTHORS     Niimi,T., Nagashima,K., Ward,J.M., Minoo,P., Zimonjic,D.B.,
            Popescu,N.C. and Kimura,S.
TITLE       Claudin-18, a novel downstream target gene for the T/EBP/NKX2.1
            homeodomain transcription factor, encodes lung- and
            stomach-specific isoforms through alternative splicing
JOURNAL     Mol. Cell. Biol. 21 (21), 7380-7390 (2001)
PUBMED      11585919
REFERENCE   2 (bases 1 to 796)
AUTHORS     Niimi,T. and Kimura,S.
TITLE       Direct Submission
JOURNAL     Submitted (15-FEB-2001) Laboratory of Metabolism, National Cancer
            Institute, NIH, 9000 Rockville Pike, Bethesda, MD 20892, USA
FEATURES    source
            location/Qualifiers
            1..796
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            10..795
            /note="alternatively spliced"

CDS

```

```
/codon_start=1
/product="claudin-18A2.1"
/protein_id="AAL15637.1"
/db_xref="GI:16224169"
/translation="MAVTACOGIGFVWSLIGIAGIIAATCMDQWSTQDLYNNPVTAVF
NYQGLRSCVRESSGFTGCRGYFTLIGLPAQVRLMIVGILGALGLISFALK
CIRISMEDSAKANWTLISGIMFIYVSGLCATAGVSFANMLVTNPMWSTANMYTCMGG
MVTQVTRITFGAALFVGVAGGLHIGVWMCIAICRGLAPETRYKAVSHASHSV
AYKPGGFKASTGFGSNTKNIYDGGARTDEVQSYPSKHDYV"

ORIGIN
Alignment Scores:
Pred. No.: 5,24e-117 Length: 796
Score: 1264.00 Matches: 240
Percent Similarity: 96.55% Conservative: 12
Best Local Similarity: 91.95% Mismatches: 9
Query Match: 93.15% Indels: 0
DB: 9 Gaps: 0

US-10-063-731-118 (1-261) x AF349452 (1-796)
QY 1 MetSerThrThrCysGlnValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
DB 10 ATGGCGGTGACTGCTGCTCAGGCGCTTGGGGTTCGTGGTTTCACTGATTGGGATGGCGC 69
QY 21 CysIleAlaAlaThrClyMetAspMetTTPSerThrGlnAspLeuTyrAspAsnProVal 40
DB 70 ATCATTTGTCGACCTGCATGACCAAGTGGAGCACCACCAAGCTTTGACAAACCCCGTA 129
QY 41 ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe 60
DB 130 ACAGCTGTTTTCACTACAGGGCTGTGGCGCTCTCTGTGTCGAGAGAGCTCTGGCTTC 189
QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
DB 190 ACCGAGTGGCGGGCTACTTCCCTGCTGGGGCTGCCAGCCATGCTGCAGGCGAGTGGCA 249
QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuValSerIlePheAla 100
DB 250 GCCCTGATGATCGTAGGCATCGCTCGGGTGCATTTGGCTCTCGTATTCATCTTTGGC 309
QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
DB 310 CTGGAATGATCGCATTTGGCAGCATGGAGGACTCTGCCAAGCCCAACATGACCTGACC 369
QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
DB 370 TCCGGGATCATGTTTCTTCTCAGGCTCTTGTGCAATTGCTGGAGTGTCTGTGTTTGGC 429
QY 141 AsnMetLeuValThrAsnPheTTPMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
DB 430 AACATCTGGTGACTAACTCTCGATGTCCACAGCTAACATGTACACCGCGCATGGGTGG 489
QY 161 MetValGlnThrValGlnThrArgTyrThrPheGlyValAlaLeuPheValGlyTTPVal 180
DB 490 ATGGTGCAGACTGTTTACAGCCAGGTACACATTTGGTGGGCTCTGTTCGTGGGCTGGGTC 549
QY 181 AlaGlyGlyLeuThrIleLeuGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
DB 550 GCTGGAGGCTCACACTAAATTTGGGGGTGTGATGATGTGATGTCATCGCTGCCGGGCGCTG 609
QY 201 ProGluGluThrAsnTyrLysAlaValSerTyrThrHisAlaSerGlyHisSerValAlaTyr 220
DB 610 CCAGAGAAACCAACTACAAAGCCGTTTCTTATCATGCTCTCAGCCGACAGTGTTCCTAC 669
QY 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysLysIle 240
DB 670 AACCTTGGAGGCTTCAAGGCCAGCATGCTGGCTTTGGGTCCCAACACCAAAACAGAGATA 729
QY 241 TyrAspGlyGlyAlaAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
DB 730 TACGATGGAGGTGCCCGCACAGAGGACGAGGTACAACTTATCTCTCCAGACGACGACT 789
```

```
QY 261 Val 261
DB 790 GTG 792

RESULT 9
AK098474
LOCUS
DEFINITION
ACCESSION
AK098474.1 GI:21758493
VERSION
oligo capping; fis (full insert sequence).
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Ninomiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T.,
Kodaira,H., Furuwa,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K.,
Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,
Suzuki,Y., Hata,H., Nakagawa,K., Mizuno,S., Morinaga,M.,
Kawamura,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,
Nishikawa,T., Sugiyama,A., Kawakami,B., Nagai,K., Isogai,T. and
Sugano,S.
NEO human cDNA sequencing project
TITLE
Unpublished
JOURNAL
2 (bases 1 to 3344)
AUTHORS
Sugano,S. and Suzuki,Y.
DIRECT SUBMISSION
Submitted (08-JUL-2002) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail: fildcna@ins.u-tokyo.ac.jp, Tel.81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT
NEO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction and 5'-end one pass sequencing: Institute of Medical
Science, University of Tokyo, Laboratory of Genome Structure, Human
Genome Center; 3'-end one pass sequencing: RAB; clone selection for
full insert sequencing: RAB and Helix Research Institute.
FEATURES
Location/Qualifiers
1..3344
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="STM00359"
/tissue_type="stomach mucosa"
/clone_lib="STM"
/note="cloning vector: pME18SFL3"

ORIGIN
Alignment Scores:
Pred. No.: 3,3e-116 Length: 3344
Score: 1264.00 Matches: 240
Percent Similarity: 96.55% Conservative: 12
Best Local Similarity: 91.95% Mismatches: 9
Query Match: 93.15% Indels: 0
DB: 9 Gaps: 0

US-10-063-731-118 (1-261) x AK098474 (1-3344)
QY 1 MetSerThrThrCysGlnValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
DB 54 ATGGCGGTGACTGCTGCTCAGGCGCTTGGGGTTCGTGGTTTCACTGATTGGGATGGCGC 113
QY 21 CysIleAlaAlaThrGlyMetAspMetTTPSerThrGlnAspLeuTyrAspAsnProVal 40
DB 114 ATCATTTGTCGACCTGCATGACCAAGTGGAGCACCACCAAGCTTTGACAAACCCCGTA 173
QY 41 ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe 60
DB 174 ACAGCTGTTTTCACTACAGGGGCTGTGGCGCTCTCTGTGTCGAGAGAGCTCTGGCTTC 233
```


Db	788	AAAGAAGATCTACGATGGGGTGGCCGACAGAACGATGAACAGTCTCATCTTACCAAG	847
Qy	258	HisAspTyrVal 261	
Db	848	TATGACTATGTG 859	
RESULT 11			
AF349450			
LOCUS			
DEFINITION		1409 bp mRNA linear ROD 18-OCT-2001	
ACCESSION		Mus musculus claudin-18A.2 mRNA, complete cds, alternatively	
VERSION		spliced.	
KEYWORDS		AF349450	
SOURCE		AF349450.1 GI:162224162	
ORGANISM		Mus musculus (house mouse)	
REFERENCE			
AUTHORS		Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Rutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus 1 (bases 1 to 1409)	
TITLE		Niimi,T., Nagashima,K., Ward,J.M., Minoo,P., Zimonjic,D.B., Popescu,N.C. and Kimura,S.	
JOURNAL		claudin-18, a novel downstream target gene for the T/EBP/NKX2.1	
MEDLINE		homeodomain transcription factor, encodes lung- and	
PUBMED		stomach-specific isoforms through alternative splicing	
AUTHORS		Mol. Cell. Biol. 21 (21), 7380-7390 (2001)	
TITLE		Submitted (15-FEB-2001) Laboratory of Metabolism, National Cancer	
JOURNAL		Institute, NIH, 9000 Rockville Pike, Bethesda, MD 20892, USA	
FEATURES			
source		Location/Qualifiers	
		1..1409	
		/organism="Mus musculus"	
		/mol_type="mRNA"	
		/db_xref="taxon:10090"	
		/notes="isolated from mixed background of 129/Sv, C57BL/6,	
		and Black Swiss mice"	
		62..688	
		/notes="alternatively spliced"	
		/codon_start=1	
		/product="claudin-18A1.2"	
		/protein_id="AAU15635.1"	
		/db_xref="GI:16224163"	
		/translation="MATTTQVVGLLSILGLACCIATGMDMWSQTDLVDNPVTAFV	
		QYGLMRSCQKSGSFFTECPYFTILGIPAWLQAVRALMIVGIVGILVLSIFALK	
		CIRIGMDDSAQKMTGTLGILFIISGICAIIVSVFANMLVTINMTANMYSGMGGK	
		MGGMVQTVQTRYTFGAALFVGWVAGGTLTGVMNCIACRGLTPDDSK"	
CDS			
Alignment Scores:			
Pred. No.:	4.83e-112	Length:	1409
Score:	1217.50	Matches:	231
Percent Similarity:	94.03%	Conservative:	21
Best Local Similarity:	86.19%	Mismatches:	9
Query Match:	89.72%	Indels:	7
DB:	10	Gaps:	2
US-10-063-731-118 (1-261) x AF349450 (1-1409)			
Qy	1	MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly	20
Db	62	ATGCCACACACACGTCGAGGTGGTGGGGTCTCTCTGTCCTCTGGGGTCTGGCGGC	121
Qy	21	CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal	40
Db	122	TGCATAGCCGCCACTGGGATGGACATGTGGAGCACTCAAGACCTGTATGACAAACCCGATC	181
Qy	41	ThrSerValPheGlnTyrGluGlyLeuTyrPargSerCysValArgGlnSerSerGlyPhe	60
Db	182	ACCOCCTGTTCCAGTATGAGGGCTCTGAGGAGTGTCTGTCACACAGATCGGGGTTTC	241

```
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/notes="isolated from mixed background of 129/Sv, C57BL/6,
and Black Swiss mice"
54.848
/notes="alternatively spliced"
/codon_start=1
/product="claudin-18A2.1"
/protein_id="AALI5636.1"
/db_xref="GI:16224166"
/translation="MSVTACQGLGVVSLIGFAGIIAATCMQDSTQDLNNPVTAVP
NYQGLWRSVCVRESSEGFTECRGYFTLLGLPAMLOALVIMVIGLVLGILVSLFALK
CIRIGSDMSAKAKMTLTSGLIFISGICALIGVSFANMLVTNFMWSTANMYSMGSG
MGWVQVQTRYTFGAALFVGWVAGGLTLGGVNMCLACRGLTPDDSNFKAHSVHSG
QNVAIRPGGPKASTGFGSNTENKKIYDGGARTEDDEQSHPTKYDV"
```

CDS

```
Db 714 GTTGCTACAGCGCTCGAGGCTTTAAGGCCAGCACTGGCTTTGGTCCACACCAAGAAC 773
Qy 238 LysLysIleTyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLys 257
Db 774 AAGAAGATCTACGATCGGGGTGCCGCCACAGAACGATGACATCTCATCTTACCACAG 833
Qy 258 HisAspTyrVal 261
Db 834 TATGACTATGTG 845
```

RESULT 13

AF349453 Mus musculus claudin-18A2.2 mRNA 861 bp linear ROD 18-OCT-2001

LOCUS AF349453 Mus musculus claudin-18A2.2 mRNA, complete cds, alternatively

spliced.

ACCESSION AF349453

VERSION AF349453.1 GI:16224171

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 861)

Authors Niimi,T., Nagashima,K., Ward,J.M., Minoo,P., Zimonjic,D.B.,

Popescu,N.C. and Kimura,S.

claudin-18, a novel downstream target gene for the T/EBP/NKX2.1

homeodomain transcription factor, encodes lung- and

stomach-specific isoforms through alternative splicing

Mol. Cell. Biol. 21 (21), 7380-7390 (2001)

PUBMED 11585919

REFERENCE 2 (bases 1 to 861)

Authors Niimi,T. and Kimura,S.

Direct Submission

TITLE Submitted (15-FEB-2001) Laboratory of Metabolism, National Cancer

Institute, NIH, 9000 Rockville Pike, Bethesda, MD 20892, USA

FEATURES

Location/Qualifiers

1..861

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/note="isolated from mixed background of 129/Sv, C57BL/6,

and Black Swiss mice"

54..680

/note="alternatively spliced"

/codon_start=1

/product="claudin-18A2.2"

/protein_id="AALI5636.1"

/db_xref="GI:16224172"

/translation="MSVTACQGLGVVSLIGFAGIIAATCMQDSTQDLNNPVTAVP

NYQGLWRSVCVRESSEGFTECRGYFTLLGLPAMLOALVIMVIGLVLGILVSLFALK

CIRIGSDMSAKAKMTLTSGLIFISGICALIGVSFANMLVTNFMWSTANMYSMGSG

MGWVQVQTRYTFGAALFVGWVAGGLTLGGVNMCLACRGLTPDDSK"

ORIGIN

Alignment Scores:
Pred. No.: 2,54e-105 Length: 861
Score: 1147.50 Matches: 217
Percent Similarity: 91.04% Conservative: 27
Best Local Similarity: 80.97% Mismatches: 17
Query Match: 84.56% Indels: 7
DB: 10 Gaps: 2

US-10-063-731-118 (1-261) x AF349453 (1-861)

```
Qy 1 MetSerThrThrCysGlnValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 54 ATGTCGGTGACCGCTCGCGGCTTGGGTGTCACATGATCGGCTTTCGGGCG 113
Qy 21 CysIleAlaAlaThrGlyMetAspMetTyrSerThrGlnAspLeuTyrAspProVal 40
Db 114 ATCATTTGACCACTTGTATGGACCATGGAGGACCCAGGATTATACACACCCGGTG 173
```

```
QY 41 ThrSerValPheGlnTyrCluGlyLeuThrArgSerCysValArgGlnSerSerGlyPhe 60
Db 174 ACCGTGTATCAACTACCAAGGGCTATGGCTTCATGCTCCGAGAGAGCTCTGGCTTC 233
QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 234 ACCGAGTCCGAGGCTACTTCACCTGTGGGGTCCAGCCATGCTGCAAGCTGTACGA 293
QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
Db 294 GGCCTGATCGTGGCATGTTCTTGGGGGTATCGGTATCGTGTCCATCTTCGCC 353
QY 101 LeuLysCysValleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 354 CTGAAGTGCATTCGCATTCGTAGCATGATCTGCCAAGCCCAAGATGACTCTGACT 413
QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 414 TCTGGGATCTTGTTCATCATCTCCGGCATCTGTGCAATCATTTGTTGTTGCC 473
QY 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThr-----Gly 157
Db 474 AACATGCTGTCGACCAACTCTGGATGTCACAGCTAACATGTACAGCGGCATGGCGGC 533
QY 158 MetGlyGlyMetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheVal 177
Db 534 ATGGGTGGCATGGTGACAGCCGTTACAGACAGGTACACCTTCGTCGATCGCTGCTG 593
QY 178 GlyTrpValAlaGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArg 197
Db 594 GGCCTGGTGTCTGGAGCCCTCACCTGATGGGGAGTGTATGATGTCATCGCTGCCGT 653
QY 198 GlyLeuAlaProGluGluThrAsn-----TyrLysAlaValSerTyrHisAla 213
Db 654 GGCCTGACACCATGACAGCAAGTGAAGTCTCCCTCCCTTCAAAAGCTGTGTACCATGCC 713
QY 214 SerGlyHisSerValAlaTyrLysProGlyGlyPheLysAlaSerThrGlyPheGlySer 233
Db 714 TCTGCCAATATGTTCCTACAGCCCTCGAGGCTTTAAGCCAGCACTGGCTTTGGGTCC 773
QY 234 AsnThrLysAsnLysLysIleTyrAspGlyGlyAlaArgThrGluAspGluValGlnSer 253
Db 774 AACACCAAGAACAGAGATCTACGATGGGGTGCCCGCACAGACGATGAACAGTCT 833
QY 254 TyrProSerLysHisAspTyrVal 261
Db 834 CATCCTACCAAGTATGACTATGTG 857

RESULT 14
BD237993
LOCUS BD237993 783 bp DNA linear PAT 17-JUL-2003
DEFINITION Gastric polypeptide ZSIG28.
ACCESSION BD237993
VERSION BD237993.1 GI:33047763
KEYWORDS JP 2002524103-A/2.
SOURCE synthetic construct
ORGANISM artificial construct
REFERENCE 1 (bases 1 to 783)
AUTHORS Sheppard,P.O. and Foley,K.P.
TITLE Gastric polypeptide ZSIG28
JOURNAL Patent: JP 2002524103-A 2 06-AUG-2002;
ZYMOGENETICS INC
COMMENT OS Artificial Sequence
PN JP 2002524103-A/2
PD 06-AUG-2002
PE 14-SEP-1999 JP 2000570197
PR 16-SEP-1998 US 09/154444
PI PAUL O SHEPPARD, KEVIN P FOLEY
PC C12N15/09,A61K38/00,A61K39/395,A61K39/395,A61P1/04,A61P1/14,
PC A61P3/08,
PC A61P5/50,A61P31/04,A61P31/10,A61P35/00,C07K14/47,C07K16/18, PC
```

```
C12N1/15,
PC C12N1/19,C12N1/21,C12N5/10,C12P21/02,C12P21/08,C12Q1/02 PC
,C12Q1/68,G01N33/15,
PC G01N33/50,G01N33/53,G01N33/577,C12N15/00,C12N5/00,A61K37/02 CC
Degenerate nucleotide sequence of zsig28 polypeptide CC n =
A,T,C or G Location/Qualifiers
PH Key misc feature (1). (783).
FT misc Location/Qualifiers
FEATURES
source
1..783
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
ORIGIN
Alignment Scores:
Pred. No.: 1.26e-102 Length: 783
Score: 1120.00 Matches: 211
Percent Similarity: 80.84% Conservative: 0
Best Local Similarity: 80.84% Mismatches: 50
Query Match: 82.54% Indels: 0
DB: 6 Gaps: 0
US-10-063-731-118 (1-261) x BD237993 (1-783)
QY 1 MetSerThrThrThrCysGlnValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 1 ATGWSNACNACNACNACNACNACNACNACNACNACNACNACNACNACNACNACN 60
QY 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal 40
Db 61 TGYATHGCGNCNACNGENATGGAYATGTGWSNACNACNACNACNACNACNACN 120
QY 41 ThrSerValPheGlnTyrGluGlyLeuThrArgSerCysValArgGlnSerSerGlyPhe 60
Db 121 ACNWSNGTNTTTCARTAYGARGGNTTNTGSMGNWSNTGYGTNMGNCARWSNNGNTTY 180
QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 181 ACNGARTGTMGCCNTTAYTTACNATHYNGNYNCCGNCNATGYTNCARGCNGTNMGN 240
QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
Db 241 GCNYTNATGATHTGNGNATHTGNTNGGNCNACNACNACNACNACNACNACN 300
QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 301 YTNAAARTGYATHTMGNATHGNGNWSNATGGARGAYWSNGCNAAARGCNAAATGACNYTNACN 360
QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 361 WNGGNCNATHATGTTTATHTGNTWSNGNYNTGYGNCNATHGCGNGGNTNWSNGNTTYGNCN 420
QY 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db 421 AAYATGTYTNGTNACNAAATTTTGGATGWSNACNACNACNACNACNACNACN 480
QY 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal 180
Db 481 ATGGTNCARACNCTNCARACNNGNTAYACNTTYGGGNCNACNNTTYTGTNGTNGTNGGNTN 540
QY 181 AlaGlyGlyLeuThrLeuIleGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db 541 GCNGGNGNYTNACNTTNTATHGNGGNGNTATGATGTGYATHGCTGYMNGGNYTNGCN 600
QY 201 ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
Db 601 CCNGARGARACNAAATAYAAARGCNGTNWSNTAYCAYGCNWSNGNCAYWSNGTNGCNTAY 660
QY 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysIle 240
Db 661 AARCCNGGNGGNTTAAARGCNWSNACNCGNTTYGNGWSNAAAYACNAAARAAARATH 720
```

Qy 241 TyrAspGlyAlaAa-gThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
 Db |||||
 721 TAYGAYGGGCGGCMGACGARGAYGARGTNCARWSNTAYCCNWSNARCAVGYATAY 780
 Qy 261 Val 261
 Db |||||
 781 GTN 783

RESULT 15

BC063845 2040 bp mRNA linear PRI 12-DEC-2003
 LOCUS Homo sapiens cDNA clone IMAGE:6177081, partial cds.

DEFINITION BC063845
 ACCESSION BC063845
 VERSION 1 GI:39795429

KEYWORDS
 SOURCE Homo sapiens (human)

ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS 1 (bases 1 to 2040)

1 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, F.,
 Diatchenko, K., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.C., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butlerfield, J.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2040)

Strausberg, R.

Direct Submission

Submitted (08-DEC-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. James R. Lupski

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Letticia Hsiao, Martin Krywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Lilja Prabhu, Parvaneh Saedi, Jacqueline
 Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAP Plate: 141 Row: a Column: 13

This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 7705960.

Location/Qualifiers

FEATURES

source

1. 2040
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6177081"
 /tissue_type="Peripheral Nervous System, dorsal root
 ganglion"
 /clone_lib="Lupski_dorsal_root_ganglion"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN

Alignment Scores:

Pred. No.: 7.64e-101 Length: 2040
 Score: 1107.50 Matches: 221
 Percent Similarity: 84.67% Conservative: 0
 Best Local Similarity: 84.67% Mismatches: 1
 Query Match: 81.61% Indels: 40
 DB: 9 Gaps: 1

US-10-063-731-118 (1-261) x BC063845 (1-2040)

Qy 1 MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
 Db 11 ATGTCCACCACCAATGCGAGTGGCGTTCCTCTCTCCATCTCTGGGGCTGGCGGC 70
 Qy 21 CysIleAlaAlaThrGlyMetAspMetTyrSerThrGlnAspLeuTyrAspAsnProVal 40
 Db 71 TGCATCGGGCCACCGGATGGATGATGTGGAGCACCCAGGACCTGTACGACACCCGTC 130
 Qy 41 ThrSerValPheGlnThrGlyGlyLeuTyrPargSerCysValArgGlnSerSerGlyPhe 60
 Db 131 ACCTCCGTGTTCAGTACGAAGGGCTCTGGAGGAGTGGCGTGGAGGAGTTCAGGCTC 190
 Qy 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
 Db 191 ACCGAATGACGGCCCTATTTCCATCTCTGGAGTTCAGGAGTTCAGGAGTGGCA 250
 Qy 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
 Db 251 GCCCTGATGATCGTAGGCATCGTCTGGTGGTCCATTTGGCTCTCTGGTATCTCTTTC 310
 Qy 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
 Db 311 CTGAATGTCATCCGATTTGGCAGCATGGAGGACTCTGCCAAGGCCAACATGACATGACC 370
 Qy 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
 Db 371 TCCGGGATCATGTTTCATGCTCTC----- 393
 Qy 141 AsnMetLeuValThrAsnPheTyrMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
 Db 393 ----- 393
 Qy 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTyrVal 180
 Db 394 -----AGGTACACATTTGGTGGTCTGTTCGTGGGCTGGGTC 432
 Qy 181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
 Db 433 GCTGGAGGCTCACACATTAATTTGGGGGTGTGATGATGATGATGCTGCCGGGGCTGGCA 492
 Qy 201 ProGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
 Db 493 CCAGAGAAACCAACTCAAAAGCGTTCCTTATCATGCTCGGCGCCACAGTGTTCCTAC 552
 Qy 221 LysProGlyGlyPheLysAlaSerThrGlyPheSerAsnThrLysAsnLysLysIle 240
 Db 553 AAGCTTGAGGCTTCAAGGGCCAGCAGTGGCTTTTGGTTCACACCAACCAAGAGATA 612
 Qy 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
 Db 613 TACGATGGAGGTGCGCGCACAGAGGACGAGGTACATCTTATCTCTCCAGCACGACTAT 672

Qy 261 Val 261
Db |||
673 GTG 675

Search completed: September 2, 2004, 21:14:00
Job time : 2837 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 2, 2004, 20:04:09 ; Search time 327 Seconds
(without alignments)
3390.764 Million cell updates/sec

Title: US-10-063-731-118
Perfect score: 1357
Sequence: 1 MSTRYQVAFLLSILGLAG.....DGGARTEDEVQSPKXHDYV 261

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US10063731/runat_01092004_155041_18736/app_query.fasta_1.455
-DB=N_Geneseq 257an04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10063731 @CGN 1.1 352 @runat_01092004_155041_18736 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq 29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1357	100.0	982	3 AAZ52249	Human sto
2	1357	100.0	1505	3 AAC74775	Human ORF
3	1357	100.0	1530	3 AAA09120	Human ORF
4	1357	100.0	1530	3 AAA09116	Human ORF
5	1357	100.0	1869	6 ABK81817	DNA repre
6	1357	100.0	2108	4 AAF54432	DNA encod
7	1357	100.0	2121	3 AAA37114	Human PRO
8	1357	100.0	2121	4 AAS46102	Human DNA

9	1357	100.0	2121	4 AAF92116	Human PRO
10	1357	100.0	2121	6 ABS74436	Human CDN
11	1357	100.0	2121	7 ABX78705	Human PRO
12	1357	100.0	2121	7 ACA75677	Novel hum
13	1357	100.0	2121	7 ACC87157	Human sec
14	1357	100.0	2121	7 ACC87685	Human sec
15	1357	100.0	2121	7 ACC87071	Human sec
16	1357	100.0	2121	7 ACD04244	Human sec
17	1357	100.0	2121	7 ACA69575	cDNA enco
18	1357	100.0	2121	7 ACA90420	Novel hum
19	1357	100.0	2121	7 ACC89527	Human sec
20	1357	100.0	2121	7 ACA98318	Novel hum
21	1357	100.0	2121	7 ACA93960	Human sec
22	1357	100.0	2121	7 ACD15353	Human sec
23	1357	100.0	2121	7 ACD08940	Human sec
24	1357	100.0	2121	7 ACC96860	Human sec
25	1357	100.0	2121	7 ACF15581	Human sec
26	1357	100.0	2121	7 ACA72948	Human PRO
27	1357	100.0	2121	7 ACD03120	Novel hum
28	1357	100.0	2121	7 ACD01935	Novel hum
29	1357	100.0	2121	7 ACA92127	Novel hum
30	1357	100.0	2121	7 ACA89552	cDNA enco
31	1357	100.0	2121	7 ACA73562	Human sec
32	1357	100.0	2121	7 ACA05877	Human sec
33	1357	100.0	2121	7 ACA66711	cDNA enco
34	1357	100.0	2121	7 ACA91222	Novel hum
35	1357	100.0	2121	7 ACD81599	Human CDN
36	1357	100.0	2121	7 ACF20286	Human sec
37	1357	100.0	2121	7 ACF19672	Human sec
38	1357	100.0	2121	7 ACD21960	Human sec
39	1357	100.0	2121	7 ACF13125	Human sec
40	1357	100.0	2121	7 ACD25228	Human sec
41	1357	100.0	2121	7 ACF00277	Human sec
42	1357	100.0	2121	7 ACA60421	Novel hum
43	1357	100.0	2121	7 ACA72334	Novel hum
44	1357	100.0	2121	7 ACD04858	Novel hum
45	1357	100.0	2121	7 ACD18319	Human sec

ALIGNMENTS

RESULT 1

AZ52249

ID AAZ52249 standard; DNA; 982 BP.

XX

XX

XX

DT 18-JUL-2000 (first entry)

XX

XX

DE Human stomach protein zsig28 DNA.

XX

XX

KW

KW

KW

KW

KW

KW

OS

OS

PH

FT CDS

FT

FT

FT

FT

FT

FT

FT

XX

XX

PN

XX

XX

PD

Human; stomach; zsig28 protein; chromosome 3q22.1-3q22.2; gene therapy; claudin; oligodendrocyte-specific protein; OSP; apoptosis; RVP.1; rat androgen-withdrawal apoptosis protein; growth factor receptor; cell-cell signalling molecule; cytostatic; antibacterial; food poisoning; Botulism; diarrhoea; inflammation; cramping; cancer; gastric ulcer; diagnosis; prevention; treatment; ds.

Homo sapiens.

Key Location/Qualifiers

CDS 70..855

/*tag= a

/product= "zsig28 protein"

FT sig_peptide

70..138

/*tag= b

139..852

FT mat_peptide

/*tag= c

/product= "Mature zsig28"

WO200015659-A2.

23-MAR-2000.

XX Shimkets RA, Leach M;
 PI WPI; 2000-602362/57.
 DR P-PSDB; AAB40566.
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease.
 XX Claim 5; Page 771; 5507pp; English.
 XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORF open reading frames 1 to 3161. The ORF
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antipsoriatic; antiparkinsonian; nontropic; neuroprotective; osteopathic;
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
 CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
 CC sequences can be used for determining the presence of or predisposition
 CC to, or preventing or treating pathological conditions associated with an
 CC ORF-associated disorder. The nucleic acids can be used to express ORF
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancers, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
 XX SQ Sequence 1505 BP; 367 A; 398 C; 337 G; 403 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6,34e-149 Length: 1505
 Score: 1357.00 Matches: 261
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-10-063-731-118 (1-261) x AAC74775 (1-1505)

QY 1 MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
 DB 10 ATGTCCACCACCATGCCAAGTGTGGGTTCCTCTGTCATCTGGGGCTGGCGGC 69
 QY 21 CysIleAlaAlaThrGlyMetAspMetTyrSerThrGlnAspLeuTyrAspAsnProVal 40
 DB 70 TGCATCGCGGCCACCGGGATGGACATGTGGAGCACCAGACCTGTACACACCCCGTC 129
 QY 41 ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe 60
 DB 130 ACTCCGTGTTCAGTACGAGGGCTCTGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTC 189
 QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
 DB 190 ACCGAATGAGGCCCTTATTCACCATCTGGGACTTCCAGCCATGTGAGGAGTGGCA 249
 QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
 DB 250 GCCTCATGATGTAGCATGTCTGGGTGCCATTTGGCTCTGATTCATCTTTGGC 309
 QY 101 LeuLeuCysIleArgIleGlySerMetGluAspSerAlaIleAlaAsnMetThrLeuThr 120
 DB 310 CTGAATGATCGCATTTGGCAGCATGGAGGACTCTGCCAAGCCCAACATGACACTGACC 369
 QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
 DB 370 TCCGGATCATGTTCATTTCTCAGGTCTTTGTGCAATTTCTGGAGTGTCTGTGTTGCC 429

QY 141 AsnMetLeuValThrAsnPheTyrMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
 DB 430 AACATGCTGGTGAATACTTCTGGATGTCCACAGCTAACATGTACACCGCATGGGTGGG 489
 QY 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTyrVal 180
 DB 490 ATGGTGCAGACTGTTTCCAGACAGGTACACATTTGGTGGGCTCTGTTCGTGGGCTGGGTC 549
 QY 181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
 DB 550 GCTGGAGGCTCACACTAATTTGGGGTGTGATGTGCATCGCTTCCGGGGCTGGCA 609
 QY 201 ProGluGlnThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
 DB 610 CCAGAGAAACCAACTACAAAGCGTTTCTTATCATGCTCAGGCCACAGTGTTCCTAC 669
 QY 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysLysIle 240
 DB 670 AAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCACACCAACCAAGAGATA 729
 QY 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
 DB 730 TAGATGGAGGTGCCCGCACAGAGGACGAGGTACAACTTATCTTCCCAAGCAGACTAT 789
 QY 261 Val 261
 DB 790 GTG 792
 XX RESULT 3
 XX AAA09120/c
 XX ID AAA09120 standard; cDNA; 1530 BP.
 XX AC AAA09120;
 XX DT 10-AUG-2000 (first entry)
 XX DE Clone 3224646 cDNA complement.
 XX KW Clone 3224646; claudin; homologue; cytostatic; anti-HIV;
 XX KW immunosuppressive; anti-allergic; anti-infective; anti-inflammatory;
 XX KW anti-arthritic; anti-arteriosclerotic; vasotropic; neuroprotective;
 XX KW nontropic; dermatological; tranquilizer; vulnery; ss.
 XX OS Homo sapiens.
 XX PN WO2000020447-A2.
 XX PD 13-APR-2000.
 XX PF 06-OCT-1999; 99WO-US023294.
 XX PR 06-OCT-1998; 98US-0103195P.
 XX PR 05-OCT-1999; 99US-00412231.
 XX PA (CURA-) CURAGEN CORP.
 XX PI Shimkets RA;
 XX DR WPI; 2000-303741/26.
 XX PT Nucleic acids encoding polypeptides with syncline-like, claudin-like or
 PT cytokine-like activity, useful for treating diseases including cancer,
 PT Alzheimer's and atherosclerosis.
 XX PS Claim 12; Fig 3B; 118pp; English.
 XX CC Clone 3223867 encodes a polypeptide that has homology to claudin-1, which
 CC is an integral membrane protein found in tight junctions. The sequences
 CC are useful for treatment of diseases such as cancer, immune disorders,
 CC autoimmune disease, acquired immune deficiency syndrome (AIDS),
 CC transplant rejection, allergy, infection by a pathological agent or
 CC organism, inflammatory disorders, arthritis, a haematopoietic disorder, a
 CC skin disorder, atherosclerosis, restenosis, a neurological disease,

CC Alzheimer's disease, trauma, spinal cord injury and skeletal disorders
XX
SQ Sequence 1530 BP; 375 A; 381 C; 408 G; 366 T; 0 U; 0 Other;

Alignment Scores:
Pred. NO.: 6.5e-149 Length: 1530
Score: 1357.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-063-731-118 (1-261) x AAA09116 (1-1530)

Qy 1 MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 1230 ATGTCACACACATGCAAGTGGTGGGTTCTCTCTGTCATCTGGGGCTGGCGGC 1171
Qy 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal 40
Db 1170 TGCATCGCGGCCACCGGATGGACATGTGGAGCACCCAGGACCTGTACGACACCCCGTC 1111
Qy 41 ThrSerValPheGlnTyrGluGlyLeuTrpArgSerCysValArgGlnSerSerGlyPhe 60
Db 1110 ACTTCGGTGTCCAGTACGAAGGGCTCTGGAGGAGCTCGTGAGGACAGTTCAGGCTTC 1051
Qy 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 1050 ACCGAATGCAGCGCCATTATTCACCATCTGGGACTTCAGGCATGCTGCAGGATGGCA 991
Qy 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuValSerIlePheAla 100
Db 990 GCCCTGATGATCGTAGGCATGCTCTGGGTGCCATTTGGCTCTCTGATTCATCTTTGCC 931
Qy 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 930 CTGAATGCATCCGATGGCAGATGGAGACTCTGCCAAGCCACATGACATGACC 871
Qy 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 870 TCCGGGATCATGTTCAATTGCTCAGGCTTTGTGCAATTGTCGAGTCTCTGTGTTGCC 811
Qy 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db 810 AACATGCTGGTGAATACTTCTGGATGTCCAGCTAACATGTACACCGCATGGGTGGG 751
Qy 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal 180
Db 750 ATGTTGCAGACTGTTTCAGACAGGTACACATTTGGTGGGCTCTGTTCTGGGCTGGGTC 691
Qy 181 AlaGlyGlyLeuThrLeuIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db 690 GCTGGAGGCTTCACACTAAATTTGGGGTGTGATGATGTCATCGCTCGCGGGGCTGGCA 631
Qy 201 ProGluGlnThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
Db 630 CCAGAAGAAACAACTACAAAGCCGTTCTTATCATGCTCAGGCCACAGTGTGGCTAC 571
Qy 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysIle 240
Db 570 AAGCTGGAGGCTTCAGGCCAGCACTGGCTTTGGTTCACACCAACCAACAGAGATA 511
Qy 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db 510 TACGATGGAGTGGCCGCACAGAGACGAGGTACAATCTTATCTTCCCAAGCAGCTAT 451
Qy 261 Val 261
Db 450 GTG 448

RESULT 4
AAA09116
ID AAA09116 standard; cDNA; 1530 BP.

XX
AC AAA09116;
XX
DT 10-AUG-2000 (first entry)
XX
DE Clone 3224646 cDNA, encodes claudin homologue.
XX
KW Clone 3224646; claudin; homologue; cytostatic; anti-HIV;
KW immunosuppressive; anti-allergic; anti-infective; anti-inflammatory;
KW anti-arthritis; anti-arteriosclerotic; vasotropic; neuroprotective;
KW nootropic; dermatological; tranquilizer; vulnerary; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 301..1086
FT /*tag= a
FT sig_peptide 301..369
FT /*tag= b
FT mat_peptide 370..1083
FT /*tag= c
XX
PN WO200020447-A2.
XX
PD 13-APR-2000.
XX
XX 06-OCT-1999; 99WO-US023294.
XX
PR 06-OCT-1998; 98US-0103195P.
PR 05-OCT-1999; 99US-00412231.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA;
XX
DR WPI; 2000-303741/26.
DR P-PSDB; AA192235.
XX
PT Nucleic acids encoding polypeptides with syncytin-like, claudin-like or
PT cytokine-like activity, useful for treating diseases including cancer,
PT Alzheimer's and atherosclerosis.
XX
PS Claim 1; Fig 3A; 118pp; English.
XX
CC Clone 3223867 encodes a polypeptide that has homology to claudin-1, which
CC is an integral membrane protein found in tight junctions. The sequences
CC are useful for treatment of diseases such as cancer, immune disorders,
CC autoimmune disease, acquired immune deficiency syndrome (AIDS),
CC transplant rejection, allergy, infection by a pathological agent or
CC organism, inflammatory disorders, arthritis, a haematopoietic disorder, a
CC skin disorder, atherosclerosis, restenosis, a neurological disease,
CC Alzheimer's disease, trauma, spinal cord injury and skeletal disorders
XX
SQ Sequence 1530 BP; 366 A; 408 C; 381 G; 375 T; 0 U; 0 Other;

Alignment Scores:
Pred. NO.: 6.5e-149 Length: 1530
Score: 1357.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-063-731-118 (1-261) x AAA09116 (1-1530)

Qy 1 MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 301 ATGTCACACACATGCAAGTGGTGGGTTCTCTCTGTCATCTGGGGCTGGCGGC 360
Qy 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal 40
Db 361 TGCATCGCGGCCACCGGATGGACATGTGGAGCACCCAGGACCTGTACGACACCCCGTC 420

QY 41 ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe 60
 Db 421 ACCTCCGTGTTCAGTACGAGAGGCTCTGGAGGAGCTGCGTGGAGCAGAGTTCCAGGCTTC 480
 QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
 Db 481 ACCGAATGCGAGGCGCTATTTCACCATCTCTGGACTTCAGCAGCATGCTGCGAGCAGTCCGA 540
 QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
 Db 541 GCCTGATGATCGTAGCATCTGCTGGGTGCCATTTGGCTCTCGGTATCCATCTTTGCC 600
 QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
 Db 601 CTGAATATGATCCGATGCGATGCGAGCATGCGAGGACTCTGCCAAGCCACCATGACACTGACC 660
 QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
 Db 661 TCCGGGATCATGTTCAATTGCTCAGGCTTTGTGCAATTGCTGAGGTGCTGTGTTTGGC 720
 QY 141 AsnMetLeuValThrAsnPheTyrMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
 Db 721 AACATGCTGTGACTAACTCTCGATGTCCACAGCTAACATGTACACCGCATGGGTGG 780
 QY 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTyrVal 180
 Db 781 ATGGTGACAGCTGTTACAGACAGGTACACATTTGGTGGGCTCTGTTTGGTGGGCTGGTC 840
 QY 181 AlaGlyGlyLeuThrIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
 Db 841 GCTGGAGGCTCTACACTAAATTTGGGGGTGTGATGATGTGCATCGCTGCCGGGCTGGCA 900
 QY 201 ProGluGlnThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
 Db 901 CCAGAGAAACCACTACAAAGCGTCTTATCATGCTCTCAGGCCACAGTGTTCCTAC 960
 QY 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysIle 240
 Db 961 AAGCCTGGAGGCTTCAAGCCAGCAGCTGGCTTTGGGTCCAAACCAACCAAGAGATA 1020
 QY 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
 Db 1021 TAGATGGAGGTGCTCCGCGCAGAGGACGAGGTACATCTTATCTTCCAGCACGACTAT 1080
 QY 261 Val 261
 Db 1081 GTG 1083
 RESULT 5
 ABK81817
 ID ABK81817 standard; DNA; 1869 BP.
 XX
 AC ABK81817;
 AC
 DT 13-AUG-2002 (first entry)
 XX
 DE DNA representing lung specific gene #3.
 XX
 KW lung specific gene; gene therapy; vaccine; lung cancer; cancer staging;
 KW cancer monitoring; cancer diagnosis; imaging lung cancer; metastases;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200218576-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 27-AUG-2001; 2001WO-US026684.
 XX
 PR 28-AUG-2000; 2000US-0228378P.
 XX
 PA (DIAD-) DIADEXUS INC.

XX Chen S, Macina RA, Sun Y, Recipon H;
 XX WPI; 2002-434904/46.
 XX New lung specific genes and their encoded proteins, useful in gene
 PT therapy or as a vaccine for treating lung cancer, as well as for
 PT measuring metastases of lung cancer, or staging, monitoring, diagnosing
 PT or imaging lung cancer.
 XX Claim 1; Page 160-161; 206pp; English.
 PS The invention describes a new lung specific gene and it's variants. The
 CC lung specific gene proteins and genes are useful in gene therapy or as a
 CC vaccine for treating lung cancer. Lung specific genes are also useful for
 CC staging, monitoring, diagnosing or imaging lung cancer, as well as for
 CC measuring metastases of lung cancer. This sequence represents a lung
 CC specific gene described in the invention
 XX Sequence 1869 BP; 485 A; 457 C; 412 G; 515 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 8,71e-149 Length: 1869
 Score: 1357.00 Matches: 261
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-063-731-118 (1-261) x ABK81817 (1-1869)
 QY 1 MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
 Db 83 ATGTCACACACACATGCCAAGTGGTGGCTTCTCTGTCATCTCTGGGCTGGCGGC 142
 QY 21 CysIleAlaAlaThrGlyMetAspMetTyrSerThrGlnAspLeuTyrAspAsnProVal 40
 Db 143 TGCATCGCGGCCACCGGATGGACATGTGGAGCACCAGGACCTGTACGACACCCCGTC 202
 QY 41 ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe 60
 Db 203 ACCTCCGTGTTCAGTACGAGGCTCTGGAGGAGCTGCGTGGAGCAGAGTTCCAGGCTTC 262
 QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
 Db 263 ACCGAATGCGAGGCTTATTCACCATCTGGGACTTCCAGCCATGTCCAGGAGTCCGA 322
 QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
 Db 323 GCCTGATGATCGTAGCATCTGCTGGGTGCCATTTGGCTCTCTGTATCCATCTTTGCC 382
 QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
 Db 383 CTGAATGATCGCATGCGATGCGAGCATGCGAGGACTCTGCCAAGCCACCATGACACTGACC 442
 QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
 Db 443 TCCGGATCATGTTCAATTGCTCAGGCTTTGTGCAATTGCTGGAGTGTCTGTGTTTGGC 502
 QY 141 AsnMetLeuValThrAsnPheTyrMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
 Db 503 AACATGCTGTGACTAACTTCTGGATGTCCAGCATGACATGTACACCGCATGGGTGGG 562
 QY 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTyrVal 180
 Db 563 ATGGTGACAGCTGTTACAGACAGGTACACATTTGGTGGGCTCTGTTTGGGCTGGGTC 622
 QY 181 AlaGlyGlyLeuThrIleGlyGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
 Db 623 GCTGGAGGCTTACACTAAATTTGGGGGTGTGATGATGTGCATCGCTGCCGGGCTGGCA 682
 QY 201 ProGluGlnThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220

Db 683 CCAGAGAAACCAACTACAAAGCCGTTTCTTATCATGCTCAGGCCACAGTGTGCTTAC 742

Qy 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysLysLys 240

Db 743 AAGCCTGGAGGCTTCAAGGCCAGCAGCTGGCTTTGGGTCCAAACACCAACAAAGAGATA 802

Qy 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260

Db 803 TACGATGGAGTGGCCGACAGAGAGACGAGGTACAACTTATCTTCTTCCAAAGCAGCTAT 862

Qy 261 Val 261

Db 863 GTG 865

RESULT 6

AAF54432

ID AAF54432 standard; DNA; 2108 BP.

XX

AC AAF54432;

XX

DT 02-APR-2001 (first entry)

XX

DE DNA encoding protein of the invention #91.

XX

KW Secreted; transmembrane; gene therapy; ss.

XX

OS Unidentified.

XX

PN WO200078961-AL.

XX

PD 28-DEC-2000.

XX

PF 18-FEB-2000; 2000WO-US004342.

XX

PR 23-JUN-1999; 99US-0141037P.

PR 20-JUL-1999; 99US-0144758P.

PR 26-JUL-1999; 99US-0145698P.

PR 01-SEP-1999; 99WO-US020111.

PR 29-OCT-1999; 99US-0162506P.

PR 30-NOV-1999; 99WO-US028313.

PR 02-DEC-1999; 99WO-US028551.

PR 16-DEC-1999; 99WO-US030095.

PR 05-JAN-2000; 2000WO-US000219.

PR 06-JAN-2000; 2000WO-US000376.

XX

PA (GETH) GENENTECH INC.

XX

PI Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S, Gao W, Goddard A, Godowski PJ, Grimaldi CU, Gurney AL, Hillan KJ; Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK; Williams PM, Wood WI;

PI

XX

DR WPI; 2001-071395/08.

XX

PT Secreted and transmembrane proteins and nucleic acids designated PRO, useful as hybridization probes, in chromosome and gene mapping and gene therapy.

PT

XX

PS Claim 2; Fig 181; 787pp; English.

XX

CC The present invention relates to secreted and transmembrane proteins. These proteins and the DNA encoding them may be used as hybridization probes, in chromosome and gene mapping and in the generation of antisense RNA and DNA. They may also be used to generate either transgenic animals or knockout animals which are in turn useful for development and screening of therapeutically useful reagents. The nucleic acids may also be used in gene therapy

XX

SQ Sequence 2108 BP; 564 A; 514 C; 487 G; 543 T; 0 U; 0 Other;

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 4 Gaps: 0

US-10-063-731-118 (1-261) x AAF54432 (1-2108)

Qy 1 MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20

Db 90 ATGTCCACCACCAATGCAAGTGGTGGCGTTCCTCTGTGCATCTGGGGGTGGCCGCGC 149

Qy 21 CysIleAlaAlaThrGlyMetAspMetTyrSerThrGlnAspLeuTyrAspAsnProVal 40

Db 150 TGCATCGCGCCACCGGATGACATGTGGAGCAGCCAGGACCTGTAGCAACCCCGTC 209

Qy 41 ThrSerValPheGlnTyrGluGlyLeuTyrArgSerCysValArgGlnSerSerGlyPhe 60

Db 210 ACTCCCGTGTCCAGTAGCAAGGCGCTCTGGAGGAGCTCGGTGAGCAGAGTTCAGGCTTC 269

Qy 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80

Db 270 ACCGAATGCAGGCCCTATTTCACCATCTGGGACTTCCAGGCATGCTCAGGCAGTGGCA 329

Qy 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100

Db 330 GCCCTGATGATCGTAGGCATCTCTGGGTGCCATTTGGCTCTCTGTATCTATCTTTGCC 389

Qy 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120

Db 390 CTGAATGATCATCCGCATTTGGCAGCATTTGGAGGACTCTGCCAAGCCAAACATGACATGACC 449

Qy 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140

Db 450 TCCGGGATCATGTTTCATTTCTCAGGTCTTGTGCAATGTGGAGTGTCTGTGTTTGGCC 509

Qy 141 AsnMetLeuValThrAsnPheTyrMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160

Db 510 AACATGCTGGTGACTTAATCTTGGATGTCACAGCTAACATGTACCCGGCATGGTGGG 569

Qy 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTyrVal 180

Db 570 ATGTGTCAGACTGTTTCAGACAGGTACACATTTGGTGGCGCTCTGTCTGGGGCTGGGTC 629

Qy 181 AlaGlyGlyLeuThrLeuIleGlyValMetMetCysIleAlaCysArgGlyLeuAla 200

Db 630 GCTGGAGGCTCACACTAATTTGGGGGTGTGATGTGTCATGCTGCCCTGCCGGGCTGGCA 689

Qy 201 ProGluGlnThrAsnTyrIleValValSerTyrHisAlaSerGlyHisSerValAlaTyr 220

Db 690 CCAGAGAAACCAACTACAAAGCCGTTTCTTATCATGCTCAGGCCACAGTGTGCTTAC 749

Qy 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysLysLys 240

Db 750 AAGCCTGGAGGCTTCAAGGCCAGCAGCTGGCTTTGGGTCCAAACACCAACAAAGAGATA 809

Qy 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260

Db 810 TACGATGGAGTGGCCGACAGAGAGACGAGGTACAACTTATCTTCTTCCAAAGCAGCTAT 869

Qy 261 Val 261

Db 870 GTG 872

RESULT 7

AAA37114

ID AAA37114 standard; cDNA; 2121 BP.

XX

AC AAA37114;

XX

DT 08-AUG-2000 (first entry)

XX

DE Human PRO1572 (UNQ778) cDNA sequence SEQ ID NO:325.

XX

KW Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;
XX transmembrane; secretion; immunoadhesion; pharmaceutical; screening; ss.

OS Homo sapiens.

PN WO200012708-A2.

XX 09-MAR-2000.

PF 01-SEP-1999; 99WO-US020111.

XX 01-SEP-1998; 98US-0098716P.

PR 01-SEP-1998; 98US-0098749P.

PR 01-SEP-1998; 98US-0098750P.

PR 02-SEP-1998; 98US-0098803P.

PR 02-SEP-1998; 98US-0098821P.

PR 02-SEP-1998; 98US-0098843P.

PR 02-SEP-1998; 98US-0098936P.

PR 09-SEP-1998; 98US-0099596P.

PR 09-SEP-1998; 98US-0099598P.

PR 09-SEP-1998; 98US-0099602P.

PR 09-SEP-1998; 98US-0099642P.

PR 10-SEP-1998; 98US-0099741P.

PR 10-SEP-1998; 98US-0099754P.

PR 10-SEP-1998; 98US-0099763P.

PR 10-SEP-1998; 98US-0099792P.

PR 10-SEP-1998; 98US-0099808P.

PR 10-SEP-1998; 98US-0099812P.

PR 10-SEP-1998; 98US-0099815P.

PR 10-SEP-1998; 98US-0099816P.

PR 15-SEP-1998; 98US-0100385P.

PR 15-SEP-1998; 98US-0100388P.

PR 15-SEP-1998; 98US-0100390P.

PR 16-SEP-1998; 98US-0100584P.

PR 16-SEP-1998; 98US-0100627P.

PR 16-SEP-1998; 98US-0100661P.

PR 16-SEP-1998; 98US-0100662P.

PR 16-SEP-1998; 98US-0100664P.

PR 17-SEP-1998; 98US-0100683P.

PR 17-SEP-1998; 98US-0100684P.

PR 17-SEP-1998; 98US-0100710P.

PR 17-SEP-1998; 98US-0100711P.

PR 17-SEP-1998; 98US-0100919P.

PR 17-SEP-1998; 98US-0100930P.

PR 18-SEP-1998; 98US-0100848P.

PR 18-SEP-1998; 98US-0100849P.

PR 02-OCT-1998; 98US-0102965P.

PR 06-OCT-1998; 98US-0103258P.

PR 06-OCT-1998; 98US-0103449P.

PR 07-OCT-1998; 98US-0103314P.

PR 07-OCT-1998; 98US-0103315P.

PR 07-OCT-1998; 98US-0103328P.

PR 07-OCT-1998; 98US-0103395P.

PR 07-OCT-1998; 98US-0103396P.

PR 07-OCT-1998; 98US-0103401P.

PR 08-OCT-1998; 98US-0103633P.

PR 08-OCT-1998; 98US-0103678P.

PR 08-OCT-1998; 98US-0103711P.

PR 14-OCT-1998; 98US-0104257P.

PR 20-OCT-1998; 98US-0104987P.

PR 20-OCT-1998; 98US-0105000P.

PR 20-OCT-1998; 98US-0105002P.

PR 21-OCT-1998; 98US-0105104P.

PR 22-OCT-1998; 98US-0105169P.

PR 22-OCT-1998; 98US-0105266P.

PR 26-OCT-1998; 98US-0105693P.

PR 26-OCT-1998; 98US-0105694P.

PR 27-OCT-1998; 98US-0105807P.

PR 27-OCT-1998; 98US-0105881P.

PR 27-OCT-1998; 98US-0105882P.

PR 27-OCT-1998; 98US-0106062P.

PR 28-OCT-1998; 98US-0106023P.

PR 28-OCT-1998; 98US-0106030P.

PR 28-OCT-1998; 98US-0106032P.

PR 28-OCT-1998; 98US-0106033P.

PR 28-OCT-1998; 98US-0106178P.

PR 29-OCT-1998; 98US-0106248P.

PR 29-OCT-1998; 98US-0106384P.

PR 29-OCT-1998; 98US-0108500P.

PR 30-OCT-1998; 98US-0106464P.

PR 03-NOV-1998; 98US-0106856P.

PR 03-NOV-1998; 98US-0106902P.

PR 03-NOV-1998; 98US-0106905P.

PR 03-NOV-1998; 98US-0108919P.

PR 03-NOV-1998; 98US-0108932P.

PR 03-NOV-1998; 98US-0106934P.

PR 10-NOV-1998; 98US-0107783P.

PR 17-NOV-1998; 98US-0108775P.

PR 17-NOV-1998; 98US-0108779P.

PR 17-NOV-1998; 98US-0108787P.

PR 17-NOV-1998; 98US-0108788P.

PR 17-NOV-1998; 98US-0108801P.

PR 17-NOV-1998; 98US-0108802P.

PR 17-NOV-1998; 98US-0108806P.

PR 17-NOV-1998; 98US-0108807P.

PR 17-NOV-1998; 98US-0108867P.

PR 18-NOV-1998; 98US-0108925P.

PR 18-NOV-1998; 98US-0108848P.

PR 18-NOV-1998; 98US-0108849P.

PR 18-NOV-1998; 98US-0108850P.

PR 18-NOV-1998; 98US-0108851P.

PR 18-NOV-1998; 98US-0108852P.

PR 18-NOV-1998; 98US-0108858P.

PR 18-NOV-1998; 98US-0108904P.

XX (GETH) GENENTECH INC.

XX Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;

PI WPI; 2000-237871/20.

XX P-PSDB; AA99432.

XX New mammalian DNA sequences encoding transmembrane, receptor or secreted

PT PRO polypeptides, useful for screening of potential peptide or small

PT molecule inhibitors of the relevant receptor/ligand interactions.

XX Claim 2; Fig 185; 773pp; English.

XX AAA37022 to AAA37144 encode the new isolated human transmembrane,
 CC receptor or secreted PRO polypeptides given in AA99340 to AA99462. The
 CC transmembrane and receptor PRO proteins can be used for screening of
 CC potential peptide or small molecule inhibitors of the relevant
 CC receptor/ligand interactions. The polypeptides and nucleotide sequences
 CC encoding then have various industrial applications, including uses as
 CC pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent PCR
 CC primers and hybridisation probes used in the isolation of the PRO
 CC polypeptides from the present invention
 XX
 SQ Sequence 2121 BP; 569 A; 516 C; 487 G; 549 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.05e-148 Length: 2121
 Score: 1357.00 Matches: 261
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-10-063-731-118 (1-261) x AAA37114 (1-2121)

Qy 1 MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyValGly 20
 Db 90 ATGTCCACCCACCATGCAAGTGGTGGCTTCTCTCTCCATCTCTGGGGCTGGCGGC 49
 Qy 21 CysIleAlaAlaThrGlyMetAspMetTyrSerThrThrGlnAspLeuTyrAspAsnProVal 40
 Db 150 TGCATCGCGGCACCGGATGGACATGTGGAGACCCAGGACCTGTACGACAAACCCGTC 209
 Qy 41 ThrSerValPheGlnTyrGluGlyLeuTyrPargSerCysValArgGlnSerSerGlyPhe 60
 Db 210 ACCTCCGCTGTCCAGTACGAAGGCTCTGGAGGAGTGCCTGAGCAGAGTTCAGGCTTC 269
 Qy 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
 Db 270 ACCGAATGCGAGGCCCTATTTCACCATCTCTGGGATCTCCAGCCATGCTGAGGCGATGGCA 329
 Qy 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
 Db 330 GCCCTGATGATCGTAGGCATGCTCTGGTGGTCCATGGCTCTCTGGTATCCATCTTGCC 389
 Qy 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
 Db 390 CTGAATGTCATCCGCAITGGCAGCATGGAGGACTCTGCCAAGCCAAACATGACATGACC 449
 Qy 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
 Db 450 TCCGGGATCATGTTTCATGTTCTCAGGCTTTGTGCAATGCTGGAGTGTCTGTGTTGCC 509
 Qy 141 AsnMetLeuValThrAsnPheTyrMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
 Db 510 AACATGCTGGTGAATCACTTCTGATGTCACAGCTAACATGATACCGGATGGTGGG 569
 Qy 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTyrVal 180
 Db 570 ATGTGACAGCTGTTCAGCAGCATGACATTTGGTGGGCTCTGTTCTGGGGCTGGGTC 629
 Qy 181 AlaGlyGlyLeuThrLeuIleGlyValMetMetCysIleAlaCysArgGlyValAla 200
 Db 630 GCTGGAGGCTTCACACTAATGGGGGTGTGATGATGTCATGCTGGGCTGGGGCTGGCA 689
 Qy 201 ProGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
 Db 690 CCAGAAGAAACCACTACAAAGCGGTTCTTATCATGCTCAGGCCACAGTGTGGCTTAC 749
 Qy 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysIle 240
 Db 750 AAGCTGGAGCTTCAAGCGCAGCTGGCTTTGGTCCAAACCAACCAAGAGATA 809
 Qy 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260

Db 810 TACATGAGGTGCCCCGACAGAGGACGAGGTACAACTTATCTTCCAAAGCAGACTAT 869
 Qy 261 Val 261
 Db 870 GTG 872
 RESULT 8
 AAS46102
 ID AAS46102 standard; cDNA; 2121 BP.
 XX AAS46102;
 AC AAS46102;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human DNA encoding PRO polypeptide sequence #178.
 KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep; ss;
 KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
 KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
 KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder;
 KW PCR primer.
 XX
 OS Homo sapiens.
 XX
 EN WO200168848-A2.
 XX
 PD 20-SEP-2001.
 XX
 XX 28-FEB-2001; 2001WO-US006520.
 XX
 XX 01-MAR-2000; 2000WO-US005601.
 PR 02-MAR-2000; 2000WO-US005841.
 PR 03-MAR-2000; 2000US-0187202P.
 PR 06-MAR-2000; 2000US-0186968P.
 PR 14-MAR-2000; 2000US-0189320P.
 PR 14-MAR-2000; 2000US-0189328P.
 PR 15-MAR-2000; 2000WO-US006884.
 PR 21-MAR-2000; 2000US-0190828P.
 PR 21-MAR-2000; 2000US-0191007P.
 PR 21-MAR-2000; 2000US-0191048P.
 PR 21-MAR-2000; 2000US-0191314P.
 PR 28-MAR-2000; 2000US-0192655P.
 PR 29-MAR-2000; 2000US-0193032P.
 PR 29-MAR-2000; 2000US-0193053P.
 PR 30-MAR-2000; 2000WO-US008439.
 PR 04-APR-2000; 2000US-0194449P.
 PR 04-APR-2000; 2000US-0194647P.
 PR 11-APR-2000; 2000US-0195975P.
 PR 11-APR-2000; 2000US-0196000P.
 PR 11-APR-2000; 2000US-0196187P.
 PR 11-APR-2000; 2000US-0196690P.
 PR 18-APR-2000; 2000US-0198121P.
 PR 18-APR-2000; 2000US-0198585P.
 PR 25-APR-2000; 2000US-0199397P.
 PR 25-APR-2000; 2000US-0199550P.
 PR 25-APR-2000; 2000US-0199654P.
 PR 03-MAY-2000; 2000US-0201516P.
 PR 17-MAY-2000; 2000WO-US013705.
 PR 22-MAY-2000; 2000WO-US014042.
 PR 30-MAY-2000; 2000WO-US014941.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 05-JUN-2000; 2000US-0209832P.
 PR 28-JUL-2000; 2000WO-US020710.
 PR 22-AUG-2000; 2000US-00644848.
 PR 08-NOV-2000; 2000WO-US023328.
 PR 08-NOV-2000; 2000WO-US030952.
 PR 01-DEC-2000; 2000WO-US032678.
 PR 20-DEC-2000; 2000WO-US034956.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski RJ, Gurney AL;

PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX WPI; 2001-602746/68.
DR P-PSDB; AAU29201.

XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT presence of tumors, such as prostate and breast tumors, in mammals and to
PT screen for modulators of the compounds.

XX Claim 2; Fig 355; 774pp; English.

XX Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR
CC primers for PRO polypeptides of the invention. The sequences of the
CC invention can be used to detect the presence of a tumor in a mammal by
CC comparing the level of expression of a PRO polypeptide in a test sample
CC of cells from the animal and a control sample of normal cells, whereby a
CC higher level of expression in the test sample indicates the presence of a
CC tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep,
CC pigs, goats and rabbits but are preferably human. The polypeptides can be
CC used to stimulate tumour necrosis factor (TNF) alpha release from human
CC blood, when contacted with it. A specific polypeptide can be used to
CC stimulate the proliferation or differentiation of chondrocyte cells. The
CC PRO proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders

XX Sequence 2121 BP; 569 A; 516 C; 487 G; 549 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1-05e-148	Length:	2121
Score:	1357.00	Matches:	261
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-10-063-731-118 (1-261) x AAS46102 (1-2121)

QY	1	MetSerThrThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly	20
DB	90	ATGTCACCCACACATGCCAAGTGTGGGTCTCTCTGTCATCTGGGGCTGGCGGC	149
QY	21	CysIleAlaAlaThrClyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal	40
DB	150	TGCATCGCGGCCACCGGGATGACATGTGGAGCACCAGACCTGTACGACAAACCCGTC	209
QY	41	ThrSerValPheGlnTyrGluGlyLeuTrpArgSerCysValArgGlnSerSerGlyPhe	60
DB	210	ACCTCGGTGTTCCAGTACGAGGGCTCTGGAGGAGCTGGTGGGAGAGTTCAGGCTTC	269
QY	61	ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetIleuGlnAlaValArg	80
DB	270	ACCGAATGAGGCCCTTATTCACCTCTGGGACTTCCAGCCATGCTGCGAGCGAGTCGA	329
QY	81	AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuValSerIlePheAla	100
DB	330	GCCTGATGATCGTAGCATCGCTGGGTGGCATTTGGCTCTGATATCCATCTTTGCC	389
QY	101	LeuLysCysIleArgIleGlySerMetGluAspSerAlaIleAlaAsnMetThrLeuThr	120
DB	390	CTGAATGATCGCATTCGACATGGCAGCTGAGGAGCTCTGCCAAGCCAAACATGACATGAC	449
QY	121	SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla	140
DB	450	TCCGGGATCATGTTCTCTCAGGTCTTTGGCAATTTGCTGAGTGTCTGTGTTTGC	509
QY	141	AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly	160
DB	510	AACATCTGGTGAACCTCTGGATGTCCACAGCTAACATGTACACCGCATGGGTGGG	569
QY	161	MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal	180

DB	570	ATGTCACACATGTTTCAGACACAGGTACACATTTGGTGGGCTCTGTTGGTGGGCTGGGTC	629
QY	181	AlaGlyGlyLeuThrIleGlyValMetMetCysIleAlaCysArgGlyLeuAla	200
DB	630	GCTGGAGGCTCACACTAATTCGGGGTGTGATGATGATGATGATGATGATGATGATGATGAT	689
QY	201	ProGluGluThrAsnTyrIleValSerIleAlaValSerIleHisAlaSerGlyHisValAlaTyr	220
DB	690	CCAGAAGAACAACCACTACAAAGCCGTTCTTATCATGCTTCAGGCCACAGTGTGCTTAC	749
QY	221	LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrIleLysAsnLysIle	240
DB	750	AAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGTCCACACCAAAACAAGAGATA	809
QY	241	TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr	260
DB	810	TACGATGGAGGTGCCGCAACAGAGGACGAGTACATCTTATCTTCCAGGACGACTAT	869
QY	261	Val 261	
DB	870	GTG 872	
RESULT 9			
AAF92116			
ID	AAF92116 standard; cDNA; 2121 BP.		
XX	AAF92116;		
XX	15-MAY-2001 (first entry)		
DE	Human PRO1572 cDNA.		
XX	Human; PRO protein; mapping; ss.		
XX	Homo sapiens.		
XX	WO200116318-A2.		
XX	08-MAR-2001.		
XX	24-AUG-2000; 2000WO-US023328.		
PR	01-SEP-1999; 99WO-US020111.		
PR	15-SEP-1999; 99WO-US021090.		
PR	07-DEC-1999; 99US-016949SP.		
PR	09-DEC-1999; 99US-0170262P.		
PR	11-JAN-2000; 2000US-0175481P.		
PR	18-FEB-2000; 2000WO-US004341.		
PR	22-FEB-2000; 2000WO-US004414.		
PR	01-MAR-2000; 2000WO-US005601.		
PR	03-MAR-2000; 2000US-0187202P.		
PR	21-MAR-2000; 2000US-0191007P.		
PR	30-MAR-2000; 2000WO-US008439.		
PR	25-APR-2000; 2000US-0199197P.		
PR	22-MAY-2000; 2000WO-US014042.		
PR	05-JUN-2000; 2000US-0209832P.		
XX	(GETH) GENENTECH INC.		
XX	Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;		
PI	Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;		
XX	WPI; 2001-183260/18.		
DR	P-PSDB; AAB87584.		
XX	Eighty four nucleic acids encoding PRO polypeptides, useful in molecular		
PT	biology, including use as hybridization probes, and in chromosome and		
PT	gene mapping.		
XX	Claim 2; Fig 117; 278pp; English.		

24-SEP-1998; 98US-0101738P.
 24-SEP-1998; 98US-0101743P.
 24-SEP-1998; 98US-0101916P.
 30-SEP-1998; 98US-0102570P.
 06-OCT-1998; 98US-0103449P.
 08-MAR-1999; 98US-0103449P.
 14-MAY-1999; 98US-0103449P.
 02-JUN-1999; 98US-0103449P.
 01-SEP-1999; 98US-0103449P.
 15-SEP-1999; 98US-0103449P.
 15-SEP-1999; 98US-0103449P.
 18-FEB-2000; 2000US-0004341.
 18-FEB-2000; 2000US-0004342.
 22-FEB-2000; 2000US-0004414.
 01-MAR-2000; 2000US-0005601.
 30-MAR-2000; 2000US-0008439.
 22-MAY-2000; 2000US-0014042.
 02-JUN-2000; 2000US-0015264.
 23-AUG-2000; 2000US-0023522.
 24-AUG-2000; 2000US-0023328.
 10-NOV-2000; 2000US-0030873.
 01-DEC-2000; 2000US-0032378.
 20-DEC-2000; 2000US-0034956.
 28-FEB-2001; 2001US-0006520.
 01-MAR-2001; 2001US-0006666.
 30-MAY-2001; 2001US-0017443.
 01-JUN-2001; 2001US-0017800.
 20-JUN-2001; 2001US-0019692.
 29-JUN-2001; 2001US-0021066.
 09-JUL-2001; 2001US-0021735.
 (GETH) GENENTECH INC.
 Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
 Grimaldi JC, Gurney AL, Watanabe CK, Wood WI;
 WPI; 2002-731348/79.
 P-PSDB; ABG95909.
 New isolated secreted and transmembrane PRO polypeptide useful for
 modulating biological activity of a cell, or for treating sports-related
 joint problems, osteoarthritis or rheumatoid arthritis.
 Claim 2; Fig 117; 399pp; English.
 The invention relates to an isolated secreted and transmembrane PRO
 polypeptide having 80 % sequence identity to a sequence appearing as
 ABG95851-ABG95934 or their associated signal peptide, or a sequence of an
 extracellular domain of the proteins with their associated signal peptide
 or lacking its associated signal peptide. Also included are the nucleic
 acids encoding the proteins, vectors, host cells, fusion proteins and
 antibodies which specifically bind to the proteins. The proteins are
 useful for detecting a polypeptide designated as A, B, C or D in a sample
 suspected of containing A, B, C or D polypeptide, by contacting the
 sample with a polypeptide designated as E, F, G, H or I (or vice versa)
 and determining the formation of a A/E, B/F, C/G, H or I polypeptide
 conjugate in the sample, where the formation of the conjugate is
 indicative of the presence of an A, B, C or D polypeptide in the sample,
 where A is a PRO10272 polypeptide, B is a PRO20110 polypeptide, C is a
 PRO10096 polypeptide, D is a PRO19760 polypeptide, E is a PRO5801
 polypeptide, F is a PRO1 polypeptide, G is a PRO20040 polypeptide, H is a
 PRO20233 polypeptide and I is a PRO1890 polypeptide. The sample comprises
 a cell suspected of expressing the A, B, C or D polypeptide. The sample
 H or I polypeptide is labeled with a detectable label or is attached to a
 solid support. The proteins are useful for linking a bioactive molecule
 to a cell expressing a polypeptide designated as A, B, C or D or E, F, G,
 H or I. The bioactive molecule is a toxin, a radiolabel or an antibody.
 The bioactive molecule causes death of the cell. A, B, C, D, E, F, G, H,
 or I, or antibodies against them are useful for modulating a biological
 activity of a cell expressing a polypeptide designated as A, B, C or D or
 E, F, G, H, or I. The cell is killed. The proteins are useful for
 identifying agonists or antagonists, for the preparation of a medicament

CC useful in the treatment of a condition which is responsive to the
 CC proteins, as molecular weight markers for protein electrophoresis
 CC purposes, and as therapeutic agents for treating sports-related joint
 CC problems, articular cartilage defects, osteoarthritis or rheumatoid
 CC arthritis. Nucleic acids encoding the proteins are useful as
 CC hybridisation probes, in chromosome and gene mapping, in the generation
 CC of anti-sense RNA and DNA, for the preparation of the proteins, to
 CC generate transgenic or knockout animals which are useful in the
 CC development and screening of therapeutic useful reagents, for chromosome
 CC identification, and in gene therapy. The antibody is useful as a
 CC therapeutic agent, in a diagnostic assay and for affinity purification of
 CC the protein from recombinant cell culture natural sources. The present
 CC sequence encodes a novel secreted or transmembrane protein of the
 CC invention
 XX
 SQ Sequence 2121 BP; 569 A; 516 C; 487 G; 549 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.05e-148 Length: 2121
 Score: 1357.00 Matches: 261
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-063-731-118 (1-261) x ABS74436 (1-2121)
 QY 1 MetSerThrThrThrCysGlnValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
 DB 90 ATGTCCACCCACACATGCAAGTGGTGGCTGCTCCCTGCTCCATCTGGGGTGGCCGGC 149
 QY 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal 40
 DB 150 TGCATCGGGCCACCGGATGGACATGTGGAGCACCAGGACCTGTACGACAAACCCGTC 209
 QY 41 ThrSerValPheGlnTyrGluGlyLeuTrpArgSerCysValArgGlnSerSerGlyPhe 60
 DB 210 ACCTCCGTGTTCCAGTACGAGGGCTCTGGAGGAGCTGCTGGAGGAGCTGCTGGCTTTC 269
 QY 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
 DB 270 ACCGAATCAGGCCCTATTTTCCATCTGGAGCTTCAGGACCTGTCAGGACATGGGGA 329
 QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
 DB 330 GCCCTGATGATCGTAGGCATGCTCTGGTGGCTTGGCTTCTGCTATTCATCTTGGC 389
 QY 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
 DB 390 CTGMAATGCATCCGCTATGGCAGCATGGAGGACTCTGCCAAAGCCCAACATGACACTGACC 449
 QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
 DB 450 TCCGGGATCATGTTCAATGTTCTCAGGCTCTTGTGCAATGCTGGAGTGTCTGTGTTGGC 509
 QY 141 AsnMetLeuValThrAsnMetTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
 DB 510 AACATGCTGGTGAATCACTTCTGGATGTCACAGCTACATGATACCCGGCATGGTGGG 569
 QY 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal 180
 DB 570 ATGGTGCAGACTGTTTACAGCAGGATACATTTGGTGGGCTCTGCTGGCTGGCTGGGTC 629
 QY 181 AlaGlyGlyLeuThrIleLeuGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
 DB 630 GCTGAGGAGCCCTCACACTAATTTGGGGGTGTGATGTGATGTGATGTGCTGGCGGGCTGGCA 689
 QY 201 ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
 DB 690 CCAGAAGAAACCACTACAGGCCGCTTCTTATCATGTCTCAGGCCACAGTGTTCCTTAC 749
 QY 221 LysProGlyGlyPheLeuAlaSerThrGlyPheGlySerAsnThrLysAsnLysIle 240

Db 750 AAGCTGGAGGCTTCAGGCCAGCACTGGCTTGGTCCACACCAAAACAAGAGATA 809
Qy 241 TyrAspGlyGlyAlaThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db 810 TACGATGAGGTGCGCCACAGAGAGGACGAGTACAACTTATCTCTCCAGCACGACTAT 869
Qy 261 Val 261
Db 870 GTG 872
RESULT 11
ABX78705
ID ABX78705 standard; cDNA; 2121 BP.
XX AC
AC ABX78705;
XX DT
DT 15-APR-2003 (first entry)
XX DE
DE Human PRO polynucleotide #178.
XX KW
KW Human; PRO; gene; ss; cytostatic; tumour; cancer; breast; lung; stomach;
KW liver; dog; cat; cow; horse; sheep; pig; goat; rabbit; ADEPT;
KW antibody-dependent enzyme mediated prodrug therapy.
XX OS
OS Homo sapiens.
XX PN
PN US2003027272-A1.
XX PD
PD 06-FEB-2003.
XX PF
PF 21-JUN-2002; 2002US-00176492.
XX PR
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 15-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 11-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090896P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 01-JUL-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.

```

PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 04-JUL-1998; 98US-0094006P.
PR 24-JUL-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096953P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Alignment Scores:
Pred. No.: 1,058-148
Score: 1357.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Length: 2121
Matches: 261
Conservative: 0
Mismatches: 0
Indels: 0

DB: 7 Gaps: 0
US-10-063-731-118 (1-261) x ABX78705 (1-2121)
Qy 1 MetSerThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 90 ATGTCCACCACCATGCAAGTGGTGGCTTCTCTCTGTCCATCCTGGGGGTGGCCGAC 149
Qy 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrAspAsnProVal 40
Db 150 TGCATCGCGGCACCGGATGACATGTGGAGCAGCCAGGACCTGTATGACGACACCCGTC 209
Qy 41 ThrSerValPheGlnTyrGluGlyLeuTrpArgSerCysValArgGlnSerSerGlyPhe 60
Db 210 ACTCCGGTTCAGTACGAGGGCTCTGGAGGAGCTGGTGAGGACAGGTTCAGGCTTC 269
Qy 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 270 ACCGAATGCAGGCCCTATTTCACCATCCTGGAGCTTCAGGCATCTCCAGGCAGTGGCA 329
Qy 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuValSerIlePheAla 100
Db 330 GCCCTGATGATCGTAGGCATCGTCTGGGTGCGCATGTGGCTCTCTGTATTCATCTTTGCC 389
Qy 101 LeuLysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
Db 390 CTGAAATGCATCCGATTTGGCAGCATGGAGGACTCTGCCAAGCCAAACATGACACTGACC 449
Qy 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
Db 450 TCCGGGATCATGTTTCATGTCTCAGGTCTTGTGCAATTGCTGGAGTGTCTGTGTTTGGC 509
Qy 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrThrGlyMetGlyGly 160
Db 510 AACATGCTGGTGACTTACTTCTGGATGTCCACAGCTAACATGTACCCGGCATGGGTGGG 569
Qy 161 MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal 180
Db 570 ATGGTGACAGCTGTTCAGCAGGATACACATTTGGTGGGCTCTGTTCGTGGGCTGGGTC 629
Qy 181 AlaGlyGlyLeuThrLeuIleGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
Db 630 GCTGGAGGCCCTCACACTAATTTGGGGTGTGATGTGATGTGATCGCTGCCCGGGCCTGGCA 689
Qy 201 ProGluGluThrAsnTyrLysAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr 220
Db 690 CCAGAAGAAACCAACTACAAAGCCGTTTCTTATCATGCTCTCAGGCACAGTGTGCCCTAC 749
Qy 221 LysProGlyGlyPheLysAlaSerThrGlyPheGlySerAsnThrLysAsnLysLysIle 240
Db 750 AAGCTGGAGGCTTCAGGCCAGCAGCTGGCTTTGGTTCACACACCAACAAACAAAGAGATA 809
Qy 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerLysHisAspTyr 260
Db 810 TACGATGGAGGTGCCCGCACAGAGCAGGTACAATCTTATCTCTTCCAAAGCAGCACTAT 869
Qy 261 Val 261
Db 870 GTG 872

RESULT 12
ACA75677
ID ACA75677 standard; cDNA; 2121 BP.
XX ACA75677;
AC
DT 07-JUL-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO1572 cDNA.
XX
KW Human; secreted and transmembrane protein; PRO; gene therapy;
KW tumour necrosis factor-alpha release; TNF-alpha release;
KW chondrocyte proliferation; chondrocyte differentiation; tumour;

```

KW adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour; gene; ss.
XX Homo sapiens.
XX US2003032127-A1.
XX 13-FEB-2003.
XX 26-JUN-2002; 2002US-00183012.
XX 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-00631120P.
PR 24-OCT-1997; 97US-00631121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-00698017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078866P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-008579P.
PR 15-MAY-1998; 98US-0085860P.
PR 15-MAY-1998; 98US-0085862P.
PR 15-MAY-1998; 98US-0085870P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087088P.
PR 28-MAY-1998; 98US-0087288P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088126P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090890P.
PR 25-JUN-1998; 98US-0090894P.
PR 25-JUN-1998; 98US-0090959P.
PR 25-JUN-1998; 98US-0090969P.
PR 25-JUN-1998; 98US-0090969P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.

PR	21-NOV-1997;	97US-00666120P.	PR	17-JUN-1998;	98US-00895598P.
PR	24-NOV-1997;	97US-00664656P.	PR	17-JUN-1998;	98US-0089653P.
PR	24-NOV-1997;	97US-0066772P.	PR	18-JUN-1998;	98US-0089908P.
PR	11-DEC-1997;	97US-00693335P.	PR	19-JUN-1998;	98US-0089952P.
PR	12-DEC-1997;	97US-0069425P.	PR	22-JUN-1998;	98US-0090246P.
PR	17-DEC-1997;	97US-0069870P.	PR	22-JUN-1998;	98US-0090252P.
PR	18-DEC-1997;	97US-0068017P.	PR	22-JUN-1998;	98US-0090254P.
PR	10-MAR-1998;	98US-007450P.	PR	24-JUN-1998;	98US-0090429P.
PR	11-MAR-1998;	98US-0077632P.	PR	24-JUN-1998;	98US-0090435P.
PR	11-MAR-1998;	98US-0077649P.	PR	24-JUN-1998;	98US-0090444P.
PR	20-MAR-1998;	98US-0078886P.	PR	24-JUN-1998;	98US-0090461P.
PR	20-MAR-1998;	98US-0078939P.	PR	24-JUN-1998;	98US-0090535P.
PR	27-MAR-1998;	98US-0079664P.	PR	24-JUN-1998;	98US-0090540P.
PR	27-MAR-1998;	98US-0079786P.	PR	25-JUN-1998;	98US-0090676P.
PR	31-MAR-1998;	98US-0080107P.	PR	25-JUN-1998;	98US-0090678P.
PR	31-MAR-1998;	98US-0080119P.	PR	25-JUN-1998;	98US-0090688P.
PR	01-APR-1998;	98US-0080327P.	PR	25-JUN-1998;	98US-0090690P.
PR	01-APR-1998;	98US-0080333P.	PR	25-JUN-1998;	98US-0090694P.
PR	08-APR-1998;	98US-0081049P.	PR	25-JUN-1998;	98US-0090695P.
PR	08-APR-1998;	98US-0081070P.	PR	25-JUN-1998;	98US-0090696P.
PR	09-APR-1998;	98US-0081195P.	PR	26-JUN-1998;	98US-00105413.
PR	15-APR-1998;	98US-0081838P.	PR	26-JUN-1998;	98US-0090862P.
PR	21-APR-1998;	98US-0082568P.	PR	26-JUN-1998;	98US-0090863P.
PR	22-APR-1998;	98US-0082569P.	PR	26-JUN-1998;	98US-0091010P.
PR	22-APR-1998;	98US-0082704P.	PR	01-JUL-1998;	98US-0091359P.
PR	28-APR-1998;	98US-0082797P.	PR	01-JUL-1998;	98US-0091544P.
PR	28-APR-1998;	98US-0083322P.	PR	02-JUL-1998;	98US-0091478P.
PR	29-APR-1998;	98US-0083495P.	PR	02-JUL-1998;	98US-0091486P.
PR	29-APR-1998;	98US-0083496P.	PR	02-JUL-1998;	98US-0091626P.
PR	29-APR-1998;	98US-0083499P.	PR	02-JUL-1998;	98US-0091628P.
PR	29-APR-1998;	98US-0083559P.	PR	02-JUL-1998;	98US-0091632P.
PR	05-MAY-1998;	98US-0084366P.	PR	04-AUG-1998;	98US-0095282P.
PR	06-MAY-1998;	98US-0084414P.	PR	10-AUG-1998;	98US-0095998P.
PR	07-MAY-1998;	98US-0084639P.	PR	10-AUG-1998;	98US-0096012P.
PR	07-MAY-1998;	98US-0084640P.	PR	17-AUG-1998;	98US-0096757P.
PR	15-MAY-1998;	98US-0085579P.	PR	17-AUG-1998;	98US-0096766P.
PR	15-MAY-1998;	98US-0085580P.	PR	17-AUG-1998;	98US-0096867P.
PR	15-MAY-1998;	98US-0085582P.	PR	17-AUG-1998;	98US-0096891P.
PR	15-MAY-1998;	98US-0085700P.	PR	17-AUG-1998;	98US-0096897P.
PR	18-MAY-1998;	98US-0086023P.	PR	18-AUG-1998;	98US-0096949P.
PR	22-MAY-1998;	98US-0086392P.	PR	18-AUG-1998;	98US-0096959P.
PR	22-MAY-1998;	98US-0086486P.	PR	18-AUG-1998;	98US-0097022P.
PR	28-MAY-1998;	98US-0087098P.	PR	26-AUG-1998;	98US-0097952P.
PR	28-MAY-1998;	98US-0087208P.	PR	26-AUG-1998;	98US-0097954P.
PR	02-JUN-1998;	98US-0087609P.	PR	26-AUG-1998;	98US-0097955P.
PR	02-JUN-1998;	98US-0087759P.	PR	26-AUG-1998;	98US-0097971P.
PR	03-JUN-1998;	98US-0087827P.	PR	26-AUG-1998;	98US-0097974P.
PR	04-JUN-1998;	98US-0088025P.	PR	26-AUG-1998;	98US-0098014P.
PR	04-JUN-1998;	98US-0088028P.	PR	01-SEP-1998;	98US-0098716P.
PR	04-JUN-1998;	98US-0088029P.	PR	01-SEP-1998;	98US-0098723P.
PR	04-JUN-1998;	98US-0088033P.	PR	02-SEP-1998;	98US-0098803P.
PR	04-JUN-1998;	98US-0088326P.	PR	02-SEP-1998;	98US-0098821P.
PR	05-JUN-1998;	98US-0088167P.	PR	02-SEP-1998;	98US-0098843P.
PR	05-JUN-1998;	98US-0088202P.	PR	09-SEP-1998;	98US-0099602P.
PR	05-JUN-1998;	98US-0088212P.	PR	09-SEP-1998;	98US-0099741P.
PR	05-JUN-1998;	98US-0088217P.	PR	10-SEP-1998;	98US-0099741P.
PR	09-JUN-1998;	98US-0088655P.	PR	10-SEP-1998;	98US-0099754P.
PR	10-JUN-1998;	98US-0088722P.	PR	10-SEP-1998;	98US-0099763P.
PR	10-JUN-1998;	98US-0088738P.	PR	15-SEP-1998;	98US-0099812P.
PR	10-JUN-1998;	98US-0088740P.	PR	15-SEP-1998;	98US-0100388P.
PR	10-JUN-1998;	98US-0088811P.	PR	16-SEP-1998;	98US-0100662P.
PR	10-JUN-1998;	98US-0088824P.	PR	16-SEP-1998;	98US-0100664P.
PR	10-JUN-1998;	98US-0088825P.	PR	16-SEP-1998;	98US-0101751P.
PR	10-JUN-1998;	98US-0088826P.	PR	16-SEP-1998;	98US-0100683P.
PR	11-JUN-1998;	98US-0088861P.	PR	17-SEP-1998;	98US-0100684P.
PR	11-JUN-1998;	98US-0088863P.	PR	17-SEP-1998;	98US-0100919P.
PR	11-JUN-1998;	98US-0088876P.	PR	17-SEP-1998;	98US-0100930P.
PR	12-JUN-1998;	98US-0089090P.	PR	18-SEP-1998;	98US-0100849P.
PR	12-JUN-1998;	98US-0089105P.	PR	18-SEP-1998;	98US-0101014P.
PR	16-JUN-1998;	98US-0089512P.	PR	18-SEP-1998;	98US-0101068P.
PR	16-JUN-1998;	98US-0089514P.	PR	23-SEP-1998;	98US-0101471P.
PR	17-JUN-1998;	98US-0089538P.	PR	23-SEP-1998;	98US-0101472P.

PR 23-SEP-1998; 98US-0101475P.
 PR 23-SEP-1998; 98US-0101477P.
 PR 24-SEP-1998; 98US-0101738P.
 PR 24-SEP-1998; 98US-0101739P.
 PR 24-SEP-1998; 98US-0101743P.
 PR 24-SEP-1998; 98US-0101922P.
 PR 25-SEP-1998; 98US-0101786P.
 PR 25-SEP-1998; 98US-0102207P.
 PR 29-SEP-1998; 98US-0102240P.
 PR 29-SEP-1998; 98US-0102330P.
 PR 29-SEP-1998; 98US-0102331P.
 PR 30-SEP-1998; 98US-0102487P.
 PR 30-SEP-1998; 98US-0102570P.
 PR 30-SEP-1998; 98US-0102571P.
 PR 01-OCT-1998; 98US-0102684P.
 PR 01-OCT-1998; 98US-0102687P.
 PR 02-OCT-1998; 98US-0102965P.
 PR 06-OCT-1998; 98US-0103258P.
 PR 06-OCT-1998; 98US-0103449P.
 PR 07-OCT-1998; 98US-00168978.

Alignment Scores:
 Pred. No.: 1.05e-148 Length: 2121
 Score: 1357.00 Matches: 261
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 7 Gaps: 0

US-10-063-731-118 (1-261) x ACA71157 (1-2121)
 QY 1 MetSerThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
 DB 90 ATGTCCACACACATGCCAGTGTGGGTTCCTCTGTCATCTCTGGGGCTGGCCGC 149
 QY 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTyrrAspAsnProVal 40
 DB 150 TGCATCGCGGCCACCGGATGGACATGTGGAGCACCAGGACCTGTAGCACAACCCCGTC 209
 QY 41 ThrSerValPheGlnTyrrGluGlyLeuTrpArgSerCysValArgGlnSerSerGlyPhe 60
 DB 210 ACCTCCGTGTTCAGTACGAGGGCTCTGGAGGAGCTGGTGAGGACAGTTCAGGCTTC 269
 QY 61 ThrGluCysArgProTyrrPheThrIleLeuGlyLeuProAlaMetIleGlnAlaValArg 80
 DB 270 ACCGATGAGCGCCCTATTTCACATCTCTGGGACTTCAGCCATGCTGCGAGCATGGCA 329
 QY 81 AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla 100
 DB 330 GCCCTGATGATCGTAGGCATCGCTCTGGGTGCCATTGGCTCTCTGTATCCATCTTTGCC 389
 QY 101 LeuIysCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr 120
 DB 390 CTGAATGATCCGTCATTTGGCAGCATGGAGGACTCTGCCAAAGCCAAACATGACCTGACC 449
 QY 121 SerGlyIleMetPheIleValSerGlyLeuCysAlaIleAlaGlyValSerValPheAla 140
 DB 450 TCCGGATCATGTTTATGTCTCAGGTCTTTGTGCAATGCTGGAGTGTCTGTGTTGCC 509
 QY 141 AsnMetLeuValThrAsnPheTrpMetSerThrAlaAsnMetTyrrThrGlyMetGlyGly 160
 DB 510 AACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGTGGG 569
 QY 161 MetValGlnThrValGlnThrArgTyrrPheGlyAlaAlaLeuPheValGlyTrpVal 180
 DB 570 ATGGTGCAAGCTGTTTCAGACAGGATACACATTTGGTGGGCTCTGTTCGTGGGCTGGGTC 629
 QY 181 AlaGlyGlyLeuThrLeuIleGlyValMetMetCysIleAlaCysArgGlyLeuAla 200
 DB 630 GCTGAGGCGCTCACACTAATCGGGGTGTGATGTGATGTGATGTCATCGCTGCCGGGCGCTGGCA 689
 QY 201 ProGluGluThrAsnTyrrIysAlaValSerTyrrHisAlaSerGlyHisSerValAlaTyrr 220

Db 690 CCAGAAGAAACCACTACAAAGCCGCTTTCTTATCATGCTCAGGCACACAGTGTGCTAC 749
 QY 221 LysProGlyGlyPheIysAlaSerThrGlyPheGlySerAsnThrIysAsnIysIle 240
 Db 750 AAGCCTGGAGGCTTCAAGCCAGCAGCACTGGCTTTGGTCCAAACCAAAACCAAGAGATA 809
 QY 241 TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrrProSerIysHisAspTyrr 260
 Db 810 TACGATGGAGGTGCCCGCACAGAGGACGAGGTACAACTTTATCTTCCNAGCAGCACTAT 869
 QY 261 Val 261
 Db 870 GTG 872

RESULT 14
 ACC87685
 ID ACC87685 standard; cDNA; 2121 BP.
 AC ACC87685;
 XX
 DT 05-AUG-2003 (first entry)
 XX Human secreted polypeptide PRO1572-encoding cDNA, SEQ ID NO:355.
 DE Human; PRO; secreted protein; transmembrane protein;
 KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
 KW chondrocyte; proliferation; differentiation; cartilage disorder;
 KW bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
 KW adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
 KW liver; drug screening; transgenic animal; genetic analysis;
 KW antiarthritic; vulnery; gene therapy; gene; ss.
 OS Homo sapiens.
 XX
 PN US2003027278-A1.
 PD 06-FEB-2003.
 XX
 PF 21-JUN-2002; 2002US-00176987.
 XX 18-SEP-1997; 97US-0059263P.
 PR 18-SEP-1997; 97US-0059266P.
 PR 17-OCT-1997; 97US-0062250P.
 PR 21-OCT-1997; 97US-0063486P.
 PR 24-OCT-1997; 97US-0063120P.
 PR 24-OCT-1997; 97US-0063121P.
 PR 28-OCT-1997; 97US-0063540P.
 PR 28-OCT-1997; 97US-0063541P.
 PR 28-OCT-1997; 97US-0063544P.
 PR 28-OCT-1997; 97US-0063564P.
 PR 29-OCT-1997; 97US-0063734P.
 PR 31-OCT-1997; 97US-0063870P.
 PR 31-OCT-1997; 97US-0064103P.
 PR 13-NOV-1997; 97US-0065311P.
 PR 21-NOV-1997; 97US-0066120P.
 PR 24-NOV-1997; 97US-0066466P.
 PR 24-NOV-1997; 97US-0066772P.
 PR 11-DEC-1997; 97US-0069335P.
 PR 12-DEC-1997; 97US-0069425P.
 PR 17-DEC-1997; 97US-0069870P.
 PR 18-DEC-1997; 97US-0068017P.
 PR 11-MAR-1998; 98US-0077450P.
 PR 11-MAR-1998; 98US-0077632P.
 PR 11-MAR-1998; 98US-0077649P.
 PR 20-MAR-1998; 98US-0078866P.
 PR 20-MAR-1998; 98US-0078939P.
 PR 27-MAR-1998; 98US-0079664P.
 PR 27-MAR-1998; 98US-0079786P.
 PR 31-MAR-1998; 98US-0080107P.
 PR 31-MAR-1998; 98US-0080194P.
 PR 01-APR-1998; 98US-0080327P.
 PR 01-APR-1998; 98US-0080333P.
 PR 08-APR-1998; 98US-0081049P.

PR 08-APR-1998; 98US-0081070P.
 PR 09-APR-1998; 98US-0081195P.
 PR 15-APR-1998; 98US-0081838P.
 PR 21-APR-1998; 98US-0082568P.
 PR 21-APR-1998; 98US-0082569P.
 PR 22-APR-1998; 98US-0082704P.
 PR 22-APR-1998; 98US-0082797P.
 PR 28-APR-1998; 98US-0083322P.
 PR 28-APR-1998; 98US-0083495P.
 PR 29-APR-1998; 98US-0083496P.
 PR 29-APR-1998; 98US-0083499P.
 PR 29-APR-1998; 98US-0083559P.
 PR 05-MAY-1998; 98US-0084366P.
 PR 06-MAY-1998; 98US-0084414P.
 PR 07-MAY-1998; 98US-0084639P.
 PR 07-MAY-1998; 98US-0084640P.
 PR 07-MAY-1998; 98US-0084643P.
 PR 15-MAY-1998; 98US-0085579P.
 PR 15-MAY-1998; 98US-0085580P.
 PR 15-MAY-1998; 98US-0085582P.
 PR 15-MAY-1998; 98US-0085700P.
 PR 18-MAY-1998; 98US-0086023P.
 PR 22-MAY-1998; 98US-0086392P.
 PR 22-MAY-1998; 98US-0086486P.
 PR 28-MAY-1998; 98US-0087098P.
 PR 28-MAY-1998; 98US-0087208P.
 PR 02-JUN-1998; 98US-0087609P.
 PR 02-JUN-1998; 98US-0087759P.
 PR 03-JUN-1998; 98US-0087827P.
 PR 04-JUN-1998; 98US-0088025P.
 PR 04-JUN-1998; 98US-0088028P.
 PR 04-JUN-1998; 98US-0088029P.
 PR 04-JUN-1998; 98US-0088033P.
 PR 04-JUN-1998; 98US-0088326P.
 PR 05-JUN-1998; 98US-0088157P.
 PR 05-JUN-1998; 98US-0088202P.
 PR 05-JUN-1998; 98US-0088212P.
 PR 05-JUN-1998; 98US-0088217P.
 PR 09-JUN-1998; 98US-0088655P.
 PR 10-JUN-1998; 98US-0088722P.
 PR 10-JUN-1998; 98US-0088738P.
 PR 10-JUN-1998; 98US-0088740P.
 PR 10-JUN-1998; 98US-0088811P.
 PR 10-JUN-1998; 98US-0088824P.
 PR 10-JUN-1998; 98US-0088825P.
 PR 10-JUN-1998; 98US-0088826P.
 PR 11-JUN-1998; 98US-0088861P.
 PR 11-JUN-1998; 98US-0088863P.
 PR 11-JUN-1998; 98US-0088876P.
 PR 12-JUN-1998; 98US-0089090P.
 PR 12-JUN-1998; 98US-0089105P.
 PR 16-JUN-1998; 98US-0089512P.
 PR 16-JUN-1998; 98US-0089514P.
 PR 17-JUN-1998; 98US-0089538P.
 PR 17-JUN-1998; 98US-0089598P.
 PR 17-JUN-1998; 98US-0089653P.
 PR 18-JUN-1998; 98US-0089908P.
 PR 19-JUN-1998; 98US-0089952P.
 PR 22-JUN-1998; 98US-0090246P.
 PR 22-JUN-1998; 98US-0090252P.
 PR 22-JUN-1998; 98US-0090254P.
 PR 24-JUN-1998; 98US-0090429P.
 PR 24-JUN-1998; 98US-0090435P.
 PR 24-JUN-1998; 98US-0090444P.
 PR 24-JUN-1998; 98US-0090461P.
 PR 24-JUN-1998; 98US-0090535P.
 PR 24-JUN-1998; 98US-0090540P.
 PR 25-JUN-1998; 98US-0090676P.
 PR 25-JUN-1998; 98US-0090678P.
 PR 25-JUN-1998; 98US-0090688P.
 PR 25-JUN-1998; 98US-0090690P.
 PR 25-JUN-1998; 98US-0090694P.
 PR 25-JUN-1998; 98US-0090695P.

PR 25-JUN-1998; 98US-0090696P.
 PR 26-JUN-1998; 98US-00105413.
 PR 26-JUN-1998; 98US-0090862P.
 PR 26-JUN-1998; 98US-0090863P.
 PR 26-JUN-1998; 98US-0091010P.
 PR 01-JUL-1998; 98US-0091359P.
 PR 01-JUL-1998; 98US-0091544P.
 PR 02-JUL-1998; 98US-0091478P.
 PR 02-JUL-1998; 98US-0091486P.
 PR 02-JUL-1998; 98US-0091626P.
 PR 02-JUL-1998; 98US-0091628P.
 PR 02-JUL-1998; 98US-0091632P.
 PR 24-JUL-1998; 98US-0094006P.
 PR 04-AUG-1998; 98US-0095282P.
 PR 10-AUG-1998; 98US-0095998P.
 PR 10-AUG-1998; 98US-0096012P.
 PR 17-AUG-1998; 98US-0096757P.
 PR 17-AUG-1998; 98US-0096766P.
 PR 17-AUG-1998; 98US-0096867P.
 PR 17-AUG-1998; 98US-0096891P.
 PR 17-AUG-1998; 98US-0096897P.
 PR 18-AUG-1998; 98US-0096949P.
 PR 18-AUG-1998; 98US-0096959P.
 PR 18-AUG-1998; 98US-0097022P.
 PR 26-AUG-1998; 98US-0097952P.
 PR 26-AUG-1998; 98US-0097954P.
 PR 26-AUG-1998; 98US-0097955P.
 PR 26-AUG-1998; 98US-0097971P.
 PR 26-AUG-1998; 98US-0097974P.
 PR 26-AUG-1998; 98US-0098014P.
 PR 01-SEP-1998; 98US-0098716P.
 PR 01-SEP-1998; 98US-0098723P.
 PR 02-SEP-1998; 98US-0098803P.
 PR 02-SEP-1998; 98US-0098821P.
 PR 02-SEP-1998; 98US-0098843P.
 PR 09-SEP-1998; 98US-0099602P.
 PR 10-SEP-1998; 98US-0099741P.
 PR 10-SEP-1998; 98US-0099754P.
 PR 10-SEP-1998; 98US-0099763P.
 PR 10-SEP-1998; 98US-0099812P.
 PR 15-SEP-1998; 98US-0100388P.
 PR 16-SEP-1998; 98US-0100662P.
 PR 16-SEP-1998; 98US-0100664P.
 PR 16-SEP-1998; 98US-0101751P.
 PR 16-SEP-1998; 98US-01019330.
 PR 17-SEP-1998; 98US-0100683P.
 PR 17-SEP-1998; 98US-0100684P.
 PR 17-SEP-1998; 98US-0100919P.
 PR 17-SEP-1998; 98US-0100930P.
 PR 18-SEP-1998; 98US-0100849P.
 PR 18-SEP-1998; 98US-0101014P.
 PR 18-SEP-1998; 98US-0101068P.
 PR 23-SEP-1998; 98US-0101471P.
 PR 23-SEP-1998; 98US-0101472P.
 PR 23-SEP-1998; 98US-0101475P.
 PR 23-SEP-1998; 98US-0101477P.
 PR 24-SEP-1998; 98US-0101738P.
 PR 24-SEP-1998; 98US-0101739P.
 PR 24-SEP-1998; 98US-0101743P.
 PR 24-SEP-1998; 98US-0101922P.
 PR 25-SEP-1998; 98US-0101786P.
 PR 29-SEP-1998; 98US-0102207P.
 PR 29-SEP-1998; 98US-0102240P.
 PR 29-SEP-1998; 98US-0102330P.
 PR 29-SEP-1998; 98US-0102331P.
 PR 30-SEP-1998; 98US-0102487P.
 PR 30-SEP-1998; 98US-0102570P.
 PR 30-SEP-1998; 98US-0102571P.
 PR 01-OCT-1998; 98US-0102684P.
 PR 01-OCT-1998; 98US-0102687P.

Alignment Scores:
 Pred. No.:

1.05e-148

Length:

2121

Score: 1357.00 Matches: 261
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 7 Gaps: 0

US-10-063-731-118 (1-261) x ACC87685 (1-2121)

Qy	1	MetSerThrThrThrCysGlnValValAlaPheLeuSerIleLeuGlyLeuAlaGly	20
Db	90	ATGTCCACCAACCAATGCTGGCGTTCCTCTGTCATCTGGGGCTGGCGGC	149
Qy	21	CysIleAlaAlaThrGlyMetAspMetTyrThrGlnAspLeuTyrAspAsnProVal	40
Db	150	TGCATCGGGCCACCGGATGGACATGTGGAGCAGCCAGGACCTGTACGACAAACCCGTC	209
Qy	41	ThrSerValPheGlnTyrGlyLeuTyrPheGlnSerCysValArgGlnSerSerGlyPhe	60
Db	210	ACCTCCGTTTCCAGTACGAAGGGCTCTGGAGGAGCTCGTGGAGCAGAGTTCAGGCTTC	269
Qy	61	ThrGlnCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg	80
Db	270	ACCAATCGAGCCCTATTCCACATCTGGAGCTTCAGGCATGCTCGAGGCGAGTGGCA	329
Qy	81	AlaLeuMetIleValGlyIleValLeuGlyAlaIleGlyLeuLeuValSerIlePheAla	100
Db	330	GCCCTGATGATCGTAGGATCGTCTGGGTGCCATTTGGCTCTGTTATCCATCTTGCC	389
Qy	101	LeuIleCysIleArgIleGlySerMetGluAspSerAlaLysAlaAsnMetThrLeuThr	120
Db	390	CTGAATGCATCCGATTTGGAGCAGTGGAGGACTCTGCCAAAGCCAAACATGACATGACC	449
Qy	121	SerGlyIleMetPheIleValSerCysIleCysAlaIleAlaGlyValSerValPheAla	140
Db	450	TCCGGGATCATGTTCAATGTTCTCAGGCTTTGTGCAATGTGGAGTGTCTGTGTTTGC	509
Qy	141	AsnMetLeuValThrAsnPheTyrMetSerThrAlaAsnMetTyrThrGlyMetGlyGly	160
Db	510	AACATGCTGGTGACTAACTCTGGATGTCACAGCTACATGATACCCGGCATGGGTGG	569
Qy	161	MetValGlnThrValGlnThrArgTyrThrPheGlyAlaAlaLeuPheValGlyTrpVal	180
Db	570	ATGGTGACAGCTGTTCCAGACAGGTACACATTTGGTGGGCTCTGTTGGTGGCTGGCTC	629
Qy	181	AlaGlyGlyLeuThrLeuIleGlyValMetMetCysIleAlaCysArgGlyLeuAla	200
Db	630	GCTGGAGCCCTCACACTAATTTGGGGGTGTGATGATGTGCATCGCTGCCGGGCGCTGGCA	689
Qy	201	ProGluGluThrAsnTyrIleAlaValSerTyrHisAlaSerGlyHisSerValAlaTyr	220
Db	690	CCAGAGAAACCACTACAAAGCCGTTTCTTATCATGCTCAGGCCACAGTGTTCCTTAC	749
Qy	221	LysProGlyGlyPheIleAlaSerThrGlyPheGlySerAsnThrLysAsnLysIle	240
Db	750	AAGCCTGGAGGCTTCAAGCCAGCAGCTGGCTTTGGGTCCAAACACCAAAACAAAGAGATA	809
Qy	241	TyrAspGlyGlyAlaArgThrGluAspGluValGlnSerTyrProSerIleHisAspTyr	260
Db	810	TACGATGAGGTGTCGACAGAGGAGGAGTACAACTTTATCTTCAAGCAGCAGTAT	869
Qy	261	Val	261
Db	870	GTG	872

RESULT 15

ACC87071

ID ACC87071 standard; cdna; 2121 BP.

XX ACC87071;

AC ACC87071;

DT 05-AUG-2003 (first entry)

DE Human secreted polypeptide PRO1572-encoding cdna, SEQ ID NO:355.

XX	Human; PRO; secreted protein; transmembrane protein;
KW	extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
KW	chondrocyte; proliferation; differentiation; cartilage disorder;
KW	bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
KW	adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
KW	liver; drug screening; transgenic animal; genetic analysis;
KW	antiarthritic; vulnery; gene therapy; gene; ss.
XX	Homo sapiens.
OS	US2003036159-A1.
PN	20-FEB-2003.
XX	02-JUL-2002; 2002US-00188773.
XX	18-SEP-1997; 97US-0059263P.
PR	18-SEP-1997; 97US-0059266P.
PR	17-OCT-1997; 97US-0062250P.
PR	21-OCT-1997; 97US-0063486P.
PR	24-OCT-1997; 97US-0063120P.
PR	28-OCT-1997; 97US-0063121P.
PR	28-OCT-1997; 97US-0063540P.
PR	28-OCT-1997; 97US-0063541P.
PR	28-OCT-1997; 97US-0063544P.
PR	28-OCT-1997; 97US-0063564P.
PR	29-OCT-1997; 97US-0063734P.
PR	31-OCT-1997; 97US-0063870P.
PR	31-OCT-1997; 97US-0064103P.
PR	13-NOV-1997; 97US-0065311P.
PR	21-NOV-1997; 97US-0066120P.
PR	24-NOV-1997; 97US-0066466P.
PR	11-DEC-1997; 97US-0066772P.
PR	17-DEC-1997; 97US-0069335P.
PR	18-DEC-1997; 97US-0069870P.
PR	10-MAR-1998; 98US-0077450P.
PR	11-MAR-1998; 98US-0077632P.
PR	11-MAR-1998; 98US-0077649P.
PR	20-MAR-1998; 98US-0078866P.
PR	27-MAR-1998; 98US-0078939P.
PR	27-MAR-1998; 98US-0079664P.
PR	31-MAR-1998; 98US-0079786P.
PR	31-MAR-1998; 98US-0080107P.
PR	01-APR-1998; 98US-0080194P.
PR	01-APR-1998; 98US-0080327P.
PR	01-APR-1998; 98US-0080333P.
PR	08-APR-1998; 98US-0081049P.
PR	09-APR-1998; 98US-0081070P.
PR	15-APR-1998; 98US-0081195P.
PR	21-APR-1998; 98US-0081838P.
PR	21-APR-1998; 98US-0082568P.
PR	22-APR-1998; 98US-0082569P.
PR	22-APR-1998; 98US-0082704P.
PR	22-APR-1998; 98US-0082797P.
PR	28-APR-1998; 98US-0083322P.
PR	29-APR-1998; 98US-0083495P.
PR	29-APR-1998; 98US-0083496P.
PR	29-APR-1998; 98US-0083499P.
PR	29-APR-1998; 98US-0083559P.
PR	06-MAY-1998; 98US-0084366P.
PR	06-MAY-1998; 98US-0084414P.
PR	07-MAY-1998; 98US-0084639P.
PR	07-MAY-1998; 98US-0084640P.
PR	15-MAY-1998; 98US-0084643P.
PR	15-MAY-1998; 98US-0085579P.
PR	15-MAY-1998; 98US-0085580P.
PR	15-MAY-1998; 98US-0085582P.
PR	15-MAY-1998; 98US-0085700P.
PR	18-MAY-1998; 98US-0086023P.
PR	22-MAY-1998; 98US-0086392P.

```
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 05-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088467P.
PR 05-JUN-1998; 98US-00888202P.
PR 05-JUN-1998; 98US-00888212P.
PR 05-JUN-1998; 98US-00888217P.
PR 09-JUN-1998; 98US-0088855P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0088909P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089658P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00103413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.

PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 13-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98MO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 25-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Alignment Scores:
Pred. No.: 1.05e-148
Score: 1357.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 7

US-10-063-731-118 (1-261) x ACC87071 (1-2121)

Qy 1 MetSerThrThrCysGlnValValAlaPheLeuLeuSerIleLeuGlyLeuAlaGly 20
Db 90 ATGTCACCAACACATGCAAGTGGCGCTTCTCTGTCATCTGGGGCTGGCGGC 149
Qy 21 CysIleAlaAlaThrGlyMetAspMetTrpSerThrGlnAspLeuTrpAspAsnProVal 40
Db 150 TGCATCGCGGCCACCGGATGGACATGTGGAGCACCCAGGACCTGTACGACCAACCCGCTC 209
Qy 41 ThrSerValPheGlnTyrGluGlyLeuTrpArgSerCysValArgGlnSerSerGlyPhe 60
Db 210 ACCTCCGTGTTCAGTACGAAGGCTCTGGAGGAGCTCGTGGAGGAGATTTCAGGCTTC 269
Qy 61 ThrGluCysArgProTyrPheThrIleLeuGlyLeuProAlaMetLeuGlnAlaValArg 80
Db 270 ACCGATGACAGGCCCTATTTCACCATCTGGGACTTCCAGCCATGTGTCAGGAGTGGCA 329
```

Search completed: September 2, 2004, 21:21:12
Job time : 335 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 14:25:29 ; Search time 3730 Seconds
(without alignments)

16980.613 Million cell updates/sec

Title: US-10-063-731-117

Perfect score: 2121

Sequence: 1 ggcctccctcaggagcgcg.....ggtaaaactaatctttaa 2121

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	786	37.1	786	29	AY402516 Homo sapi
2	785.2	37.0	1751	11	AK033657 Mus muscu
3	694.2	32.7	975	13	B038449 AGENCOURT
4	674.2	31.8	679	12	BM791434 K-EST0071

5	664	31.3	675	12	BG547464
6	656.8	31.0	743	12	BM989992
7	647.2	30.5	1138	14	CF110345
8	623	29.4	680	12	BM759274
9	607.6	28.6	795	29	AY402518
10	593.2	28.0	617	29	AY402517
11	588.2	27.7	627	12	BQ016897
12	552	26.0	743	12	BI823718
13	525.8	24.8	681	14	CB462343
14	519	24.5	651	12	BM837765
15	495.8	23.4	653	12	BM788945
16	434	20.5	549	12	BM791457
17	413.2	19.5	961	12	BG538299
18	394	18.6	567	13	BX528145
19	386	18.2	627	10	BB663682
20	378	17.8	540	12	BM836731
21	377.2	17.8	600	12	BM997852
22	374.6	17.7	797	13	BU946638
23	373.6	17.6	421	10	BF091029
24	360.8	17.0	379	10	BE841291
25	349.6	16.5	420	14	N99013
26	346.8	16.4	458	10	BE841654
27	341	16.1	341	12	BM836109
28	340.2	16.0	412	10	AW867300
29	339.6	16.0	402	12	BG937956
30	339	16.0	493	14	CA436277
31	339	16.0	552	12	BM819720
32	339	16.0	567	12	BM819706
33	338.8	16.0	570	12	BM794650
34	337.4	15.9	458	13	BX108271
35	336	15.8	395	9	AA961052
36	335.8	15.8	585	12	BM785703
37	335.8	15.8	661	12	BM791792
38	335.6	15.8	633	10	BB625263
39	335.4	15.8	403	9	AI926840
40	334.4	15.8	483	9	AI187329
41	334.2	15.8	593	12	BM829469
42	333.8	15.7	367	9	AW027675
43	335.4	14.9	649	10	B8529462
44	308.2	14.5	548	9	AA882221
45	305.8	14.5	731	29	CE210717
46	305.8	14.4	660	29	AG148436
47	301	14.2	622	10	BF426746
48	300.4	14.2	428	10	AW319069
49	295	13.9	484	9	AA028634
50	287.2	13.5	500	14	CD244706
51	283.8	13.4	441	12	BM839206
52	268.2	12.6	633	10	BF426867
53	249.2	11.7	374	13	BY063182
54	248.8	11.7	1069	14	CB596764
55	236.8	11.2	455	12	BM751562
56	231	10.9	232	14	N99040
57	219.4	10.3	262	9	AI872593
58	202	9.5	520	14	CB220415
59	188.2	8.9	550	13	BQ210042
60	183	8.6	542	28	AQ426747
61	183	8.6	740	14	CB597222
62	182.2	8.6	3688	11	BC033908
63	179.6	8.5	670	29	AG086778
64	179	8.4	449	9	AI365624
65	178.2	8.4	704	28	BZ598866
66	177.6	8.4	884	13	BU187437
67	177	8.3	660	29	AG142871
68	176.4	8.3	724	29	AG117953
69	176	8.3	424	14	H01809
70	176	8.3	461	10	AW817385
71	175.8	8.3	978	12	BM475038
72	175.6	8.3	333	9	AA730646
73	174.6	8.2	466	14	H07128
74	174.2	8.2	544	13	EX491755
75	174	8.2	774	28	AQ745937
76	173.4	8.2	662	29	AG182637
77	173.2	8.2	649	13	BU617622

78	173.2	8.2	819	29	CG734227	CG734227 RP11-224F
79	172.8	8.1	664	28	AQ491423	AQ491423 RPI1-11-2
80	172.6	8.1	358	28	AQ665179	AQ665179 HS 5344 B
81	172.4	8.1	185	10	BE937032	BE937032 PM3-SN002
82	172.2	8.1	654	29	AG042723	AG042723 Pan trogl
83	172.2	8.1	696	13	EX485530	EX485530 DKFZp8661
84	172	8.1	685	29	AG187749	AG187749 Pan trogl
85	171.8	8.1	630	28	AQ346241	AQ346241 RPI11-11
86	171.2	8.1	700	29	AG114100	AG114100 Pan trogl
87	171.2	8.1	711	9	AV376479	AV376479 AV376479
88	171.2	8.1	720	10	BF037410	BF037410 601460777
89	171	8.1	527	13	BU740001	BU740001 UI-E-E00-
90	170.2	8.0	445	28	AQ664011	AQ664011 HS 2277 B
91	170	8.0	629	28	AQ422183	AQ422183 RPI1-11-2
92	170	8.0	745	29	AG015273	AG015273 Homo sapi
93	170	8.0	750	29	AG015272	AG015272 Homo sapi
94	170	8.0	810	28	AQ780966	AQ780966 HS 3138 B
95	169.8	8.0	398	14	H75272	H75272 YU07901.r1
96	169.6	8.0	422	28	B82626	B82626 RPI11-17C1
97	169.6	8.0	486	28	AQ546155	AQ546155 CITB1-R1-
98	169.6	8.0	682	10	BF677863	BF677863 602085377
99	169.4	8.0	441	9	AA859001	AA859001 UI-R-A0-b
100	169.4	8.0	533	28	AQ626520	AQ626520 CITB1-R1-

ALIGNMENTS

RESULT 1
 AY402516 786 bp DNA linear GSS 15-DEC-2003
 LOCUS Homo sapiens CLDN18 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION AY402516
 VERSION AY402516.1 GI:39758502
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
 REFERENCE 1 (bases 1 to 786)
 Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 Science 302 (5652), 1960-1963 (2003)
 JOURNAL 14671302
 PUBMED 2 (bases 1 to 786)
 REFERENCE Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source Location/Qualifiers
 1..786
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..>786
 /gene="CLDN18"
 /locus_tag="HCM1250"

Query Match 37.1%; Score 786; DB 29; Length 786;
 Best Local Similarity 100.0%; Pred. No. 6.1e-153;
 Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	90	ATGTCCACCA	CA	CATG	CCAAAGTGGTGGCGTTCTCTCTGTCCATCTCTGGGGTGGCGGC	149
Db	1	ATGTCCACCA	CA	CATG	CCAAAGTGGTGGCGTTCTCTCTGTCCATCTCTGGGGTGGCGGC	60
QY	150	TGCATCGGGC	CA	CGGGATG	ACATGTGGAGCACCAGGACCTGTACGACAAACCCGTC	209
Db	61	TGCATCGGGC	CA	CGGGATG	ACATGTGGAGCACCAGGACCTGTACGACAAACCCGTC	120
QY	210	ACCTCCGTTT	CCAGT	ACGAAGGGCTCTGGAGGAGCTGGTGAGGCAGAGTTTCAGGCTTC	269	
Db	121	ACCTCCGTTT	CCAGT	ACGAAGGGCTCTGGAGGAGCTGGTGAGGCAGAGTTTCAGGCTTC	180	
QY	270	ACCGAATG	CAGGCCCTATT	TTCACCATCTCGGAGCTTCAGGCATCTGCAAGGCAGTGGCA	329	
Db	181	ACCGAATG	CAGGCCCTATT	TTCACCATCTCGGAGCTTCAGGCATCTGCAAGGCAGTGGCA	240	
QY	330	GCCTGTGATG	TCTAGGCATCGTCTCGGGTGCCATTTGGGCTCTCTGGTATCCATCTTTGGC	389		
Db	241	GCCTGTGATG	TCTAGGCATCGTCTCGGGTGCCATTTGGGCTCTCTGGTATCCATCTTTGGC	300		
QY	390	CTGAATGATC	CGCATTTGGCAGCATGGAGGACTCTGCCAAGCCAAACATGACATGACCTGACC	449		
Db	301	CTGAATGATC	CGCATTTGGCAGCATGGAGGACTCTGCCAAGCCAAACATGACATGACCTGACC	360		
QY	450	TCCGGGATCATG	TTCATTTGTCTCAGTCTTTTGTGCAATTTCTGGAGTGTCTGTGTTGGC	509		
Db	361	TCCGGGATCATG	TTCATTTGTCTCAGTCTTTTGTGCAATTTCTGGAGTGTCTGTGTTGGC	420		
QY	510	AACATGCTGGT	GACTAACTTTCTGGATGTCACAGCTAACATGTACACCGGCATGGGTGGG	569		
Db	421	AACATGCTGGT	GACTAACTTTCTGGATGTCACAGCTAACATGTACACCGGCATGGGTGGG	480		
QY	570	ATGTTGCAGAC	CTGTTT	CAGACAGGTACACATTTGGTGGGCTCTGTTCTGTTGGGCTGGGTC	629	
Db	481	ATGTTGCAGAC	CTGTTT	CAGACAGGTACACATTTGGTGGGCTCTGTTCTGTTGGGCTGGGTC	540	
QY	630	GCTGAGGCTC	CACACTAATTTGGGGGTGTGATGATGTGATCGCCTCGCGGGGCTCTGGCA	689		
Db	541	GCTGAGGCTC	CACACTAATTTGGGGGTGTGATGATGTGATCGCCTCGCGGGGCTCTGGCA	600		
QY	690	CCAGAAGAAA	CCAACT	CAAAAGCCGTTTCTTATCATGCTCAGGCCACAGTGTTCGCTAC	749	
Db	601	CCAGAAGAAA	CCAACT	CAAAAGCCGTTTCTTATCATGCTCAGGCCACAGTGTTCGCTAC	660	
QY	750	AAGCTTGAGG	CTTCAAGGCCAGCACTGGCTTTGGGTCCAAACCAACAAAGAGATA	809		
Db	661	AAGCTTGAGG	CTTCAAGGCCAGCACTGGCTTTGGGTCCAAACCAACAAAGAGATA	720		
QY	810	TACGATGAGG	TGCCCGCACAGAGGACGAGGTACAACTTTATCTTCCAGCAGCACTAT	869		
Db	721	TACGATGAGG	TGCCCGCACAGAGGACGAGGTACAACTTTATCTTCCAGCAGCACTAT	780		
QY	870	GTGTAA	875			
Db	781	GTGTAA	786			

RESULT 2
 AK033657
 LOCUS Mus musculus adult male cecum cDNA, RIKEN full-length enriched
 library, clone:9130209K06 product:CLAUDIN-18A2.1, full insert
 sequence.
 ACCESSION AK033657
 VERSION AK033657.1 GI:26329346
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning

771 AACAAAGATCTACGATGGGGTCCCGCAGACAGCATGAACAGCTCTCATCTACC 830
858 AAGCAGCATATGTATGCTCTAAGACCTCTCAGCAGCGCGGAGAAACTCCC----- 913
831 AAGTATGACTATGTGTAGTGTCTTAAGACCGCCCAACCTGTGTGAGGAGGAACCTTCC 890
914 ----GGAGAGCTCACCCCAAAAAAACAAGAGATCCCATCTAGATTCTCTCTGCTTTGAC 969
891 CCAAGAGAGCTCACCC--CAAGCAACGGAGTCTACCTTGTCCCTTGTGATTCAAC 949
970 TCAGCTGGAAGTTAGAAAAGCTCGATTCACTTTTGGAGAGCGCAAAATGTGCTTAGC 1029
950 TGACATCTGAAAGTTGGTAAAGCTGATTTTTCATCCATAGGAGGCTAGACAGTCTTGGC 1009
1030 CTCA-GTCTGTCTCTAAATATCCACCAATAAACAGCTCAGTTA---TTTATGAATTA 1085
1010 CACATGTCTGCTCTAAATATCCCATCAAAACAGCTGAGTTATCGTTTATGAGTTA 1069
1086 GAGCTATAGCTCAATTTTCAATCTCTATTTCTTTTAAATATAACT--TTCTACT 1143
1070 GAGCCATAACACTCACTTTAGCCCAACCTCTGCTTTTACCGTAGACTTCTTTTCAT 1129
1144 CTGATGAGAGAAATGTGGTTTAACTCTCTCTCACTTTTGTATGATTTAGACAGACTCCC 1203
1130 CTGGTATGGAATGGAATTTGACTCACAGCTAATACCTTAAATGGTTTATGAGAACTTTC 1189
1204 CCTCTCTCTCTAGTCAATAACCCATTTGATGATCTATTTCCAGCTT-ATCCCCAGAA 1262
1190 ---CTTCTCTGACTTAATAAGCTGTGATGCTGATTTTCCAGCTTGCACCAAGG 1246
1263 AACTTTTGAAGAGAGTAGACCCCAAGATGTTTATTTCTGCTGTTTGAATTTGCT 1322
1247 AAATTTTAAAGGAAAAAATAATCAATTAAGGCAATTTTCTTACTCAATTTGCGCTT 1306
1323 CCCACCCCAACTTGGC-----TAGTAATAAACACTTACTGAAGAGAGCAATAAGA 1376
1307 ACCCAACCCCAACTTGAATGATAATAATAAGAACACCACTTAAGAAGAGATGCCAGAG 1366
1377 GAAAGATATTTGT-AATCTCTCCAGCCCATGATCTCGGTTTCTTACACTGTGATCTTAA 1435
1367 GAAAGATGTTGTGTTTCCCGCCAGCAGTCACTGAGTCCCGCTATGTTGATCTAGA 1426
1436 AAGTTACAAACCAAGTCAATTTTCAGTTTGAGGCAACCAACCTTTCTACTGCTGTGA 1495
1427 ACATTA-CTCGCCACAGTGAATTTTCA-AAGAGGCAAGCGAGCTGTTGCTCTCTCAG 1484
1496 CATCTCTTATATAGCAACACCATTTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCTTC 1555
1485 CATCTGCTGATTTCCAGCAAGGCCCTTCCAGAGCTTTCCACTA-----GAAGTCTCTC 1535
1556 TTCTGTCGCGGTTCAGAAATTTGCTCCCTAGATGAATGAGAA--ATTATTTTAAAT 1612
1536 CTCTCTCGAAGTCAGAAATTTCCCTTAGAGAGTAGAATAAGATTCTTTTGGGTAA 1595
1613 TTAAGTCTTAATATAGTTAAATAAATAGTTTATGTTTAAATGATACATATCTGT 1672
1596 CTGAGTCTAGTATAGTTAATAAATAGTATATAGCAAAACGGTTTGGTATCTCAGT 1655
1673 GAAATAGCTCACCCCTACATGTGATAGAGGAATAAGAAATAATTTGCTTTGACATT 1732
1656 GAAATAGTTTTCAGCTTACATAGAAAAAGCTGGGAAAAAAGATCCCTTGACATT 1715
1733 GTCTATATGGTA 1744
1716 GTCTATAGGTA 1727

RESULT 3
BU838449
LOCUS
DEFINITION AGENCOURT 8123082 Lupski dorsal root ganglion Homo sapiens CDNA
clone IMAGE:6177081 5', mRNA sequence.
ACCESSION BU838449

VERSION BU838449.1 GI:24022844
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 975)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: csapsb-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13554 row: h column: 10
High quality sequence stop: 636.
FEATURES
Location/Qualifiers
1..975
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6177081"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupski dorsal root ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGCGCTCCG-3' and
5'-GACTAGTTCTAGATCGAGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

ORIGIN
Query Match 32.7%; Score 694.2; DB 13; Length 975;
Best Local Similarity 86.4%; Pred. No. 7.5e-134;
Matches 852; Conservative 0; Mismatches 14; Indels 120; Gaps 3;
QY 81 GCCAGGATCATGTCCACCAACATGCCAAGTGTGGTTCCTCTCTCCATCTCGGG 140
Db 1 GCCAGGATCATGTCCACCAACATGCCAAGTGTGGTTCCTCTCTCCATCTCGGG 60
QY 141 CTGCGCGCTCATCGCGCCACCGGATGACATGTGGAGCACCAGACCTGTACAC 200
Db 61 CTGCGCGCTCATCGCGCCACCGGATGACATGTGGAGCACCAGACCTGTACAC 120
QY 201 AACCCCGTCACTCCGTCTTCCAGTACCAAGGCTCTCGAGAGCTGTGGGAGAGT 260
Db 121 AACCCCGTCACTCCGTCTTCCAGTACCAAGGCTCTCGAGAGCTGTGGGAGAGT 180
QY 261 TCAGCTTTCACGAATGAGGCTTATTTTACATCTCTGGGAGCTTCAGACCATGCTG 320
Db 181 TCAGCTTTCACGAATGAGGCTTATTTTACATCTCTGGGAGCTTCAGACCATG 240
QY 321 GCAGTGGAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 380
Db 241 GCAGTGGAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 381 ATCTTTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 440
Db 301 ATCTTTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 441 ACAGTGACCTCGGGATCATGTTTTCATGCTTCAGGTCTTTTGTGCAATTTGCT 500

[illegible]

Db	181	ATGAGAGATGTGGTTTTTAATCTCTCTCACAATTTTGATGATTTAGACAGACTCCCT	240
Qy	1207	CTTCTCTCTAGTCAATAAAACCCCAATGATGATCTATTTCAGCTTATCCCAAGAAAAC	1266
Db	241	CTTCTCTCTAGTCAATAAAACCCCAATGATGATCTATTTCAGCTTATCCCAAGAAAAC	300
Qy	1267	TTTGAAGGAGAGTAGACCCCAAGATGTTATTTTCTGCTGTTTGAATTTTGTCTCCCC	1326
Db	301	TTTGAAGGAGAGTAGACCCCAAGATGTTATTTTCTGCTGTTTGAATTTTGTCTCCCC	360
Qy	1327	ACCCCAACCTTGCTAGTAAATAACACTTACTGAAGAAGCAATTAAGAGAAAGATATT	1386
Db	361	ACCCCAACCTTGCTAGTAAATAACACTTACTGAAGAAGCAATTAAGAGAAAGATATT	420
Qy	1387	TGTAATCTCTCAGAGCCATGATCTCGTTTCTTACACATGATCTTAAAGTTACCAA	1446
Db	421	TGTAATCTCTCAGAGCCATGATCTCGTTTCTTACACATGATCTTAAAGTTACCAA	480
Qy	1447	CCAAAGTCATTTTCAGTTTGAGGCAACCAACCTTTCTACTGCTGTGTGACATCTCTTAT	1506
Db	481	CCAAAGTCATTTTCAGTTTGAGGCAACCAACCTTTCTACTGCTGTGTGACATCTCTTAT	540
Qy	1507	TACAGCAACACCAATCTAGAGTTTCCTGAGCTCTCCACTGAGAGTCTCTTTCTGTGCG	1566
Db	541	TACAGCAACACCAATCTAGAGTTTCCTGAGCTCTCCACTGAGAGTCTCTTTCTGTGCG	600
Qy	1567	GGTCAGAAATGTCCTAGATGAATGAGAAATATTTTTTTTAATTAAGTCCCTAAATA	1626
Db	601	GGTCAGAAATGTCCTAGATGAATGAGAAATATTTTTTTTAATTAAGTCCCTAAATA	659
Qy	1627	TAGTTAAATAAATAA 1642	
Db	660	TAGTTAAATAAATAA 675	
RESULT 6			
BM989992			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			

QY	496	TGTCCTGTTGTTGCCAACATGCTGGTGA	CTAACTTCTGGATGTCACAGCTAA	CAATGTACA	555
Db	374	TGTCCTGTTGTTGCCAACATGCTGGTAA	CCAACTTCTGGATGTCACAGCTAA	CAATGTACA	433
QY	556	CCGGCATGGCTGGGATGGTGAGACTGTT	CAGACCAAGGTACACATTTGGTGCGCTCTGT		615
Db	434	CGCGCATGGGTTGGGATGGTGAGACTGTT	CAGACCAAGGTACACATTTGGTGCGCTCTGT		493
QY	616	TCGTGGGCTGGGTCGGTCGGAGGCTCACA	CTAAATGGGGGTGTGATGATGTGCATCGCT		675
Db	494	TTGTGGGCTGGATCGCTGGAGGCTCACA	CTAAATGGAGGTGTGATGATGTGCATCGCT		553
QY	676	GCGGGGGCTGGGACACAGAGAAACCAAC	CAACTACAAAGCGGTTCTTATCATGCGCTCAGGC		735
Db	554	GCCTGGGCTGGACTCTCTGATGACCGCA	ACTTCAAAGCGGTCTCTACACAGCTTCTGGCC		613
QY	736	ACAGTGTGTGCTACAAGCGCTGGAGGCTT	CAAGGCGCAGCACTGGGCTTTGGGTCCAAACCA		795
Db	614	AAAATGTGCTACAAGCGCTGGAGGCTT	TAAGGCCAGCACTGGGCTTTGGGTCCAAACCA		673
QY	796	AAAACAGAAAGATATAGATGGAGGTGCC	CGGCACAGAGGACGAGTACATCTTATCTTT		855
Db	674	GAACAGAAAGATATATGATGGGGGTGCC	CGGCACAGAGGACGAGTACATCTTATCTTT		733
QY	856	CCAAGCACGACTATGTCTAAATGCTCTA	AGACCTCTCAGC-----ACGGCGGAAGAAA-		908
Db	734	CCAAGTACGACTATGTGTAGTGTCTTA	AGACCGGCCCGCCCAACATGTGTGCAGAGAAGAC		793
QY	909	-----CTCCCGGAGAGCTCA	CCCCAAAACAAAGGAGATCCCATCTAGATTTCTTCTTGCTTT		964
Db	794	CTTTCCCAAGGAAAGCTCA	CCCCCAAGCAAGGAGATTTCTACCTTGTTCCTTGTGCTTT		853
QY	965	TTGACTCACAGCTGGAAAGTTAGAAAGC	CTCGATTTTCATCTTTTGAGAGGCCAAATGGTC		1024
Db	854	TTGACTCACCGTTGAAAGTTGGTAAAGT	CTTAAATTTTCATCTGTGGGAGGCTAGACAGTC		913
QY	1025	TTAGCCTTCAGTCTCTGTCTCTAAATAT	TTCACCACATAAACAGCTGAGTATTATGAAAT		1084
Db	914	TTGGCCACA-----CGTCTCTAA	NATCTCCATCACAAAACAACCGAGTTACCATTTA--C		966
QY	1085	AGAGGCTATAGCTCACATTTTCAATCT	TATTTCTTTTAAATATACCTTTCTACTC		1144
Db	967	AGGTTAAGAGACCAACACACCCACCC	CTTCCCAATCTCTGTTTTTACCATAGTAAATC		1026
QY	1145	TGATGAGAGAAATGGGTTTTAATCTCT	CTCTCACATTTTGATGATTTTAGACAGACT		1200
Db	1027	TGGTGATGGAATGTAATTCGAC	TATGAGTATCTTGGATATTTCTAGAAATTT		1082
RESULT 8					
BM759274					
LOCUS					
DEFINITION	BM759274	680 bp	mrna	linear	EST 04-MAR-2002
	K-EST0039301	S9SNU601	Homo sapiens	cdna clone	S9SNU601-22-E05 5',
					rna sequence.
ACCESSION	BM759274				
VERSION	BM759274.1	GI:19088889			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 680)				
AUTHORS	Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,				
	Kim,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and				
	Kim,Y.S.				
TITLE	21C Frontier Korean EST Project 2001				
JOURNAL	Unpublished (2002)				
COMMENT	Contact: Kim YS				
	Genome Research Center				
	Korea Research Institute of Bioscience & Biotechnology				
	52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea				
	Tel: +82-42-860-4470				

1706 AATGAAAAAATTAATTCGCTTTTGACATGCTCTATATGCTACTTTTGTAAAGTCATGCTTAAG 1766
 Db AATGAAAAAATTAATTCGCTTTTGACATGCTCTATATGCTACTTTTGTAAAGTCATGCTTAAG 600
 1766 TACAAATTCATGAAAGCTCAC 1788
 Db TACAAATTCATGAAAGCTCAC 623
 RESULT 9
 AY402518 795 bp DNA linear GSS 15-DEC-2003
 LOCUS Mus musculus CLDN18 gene, VIRTUAL TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 ACCESSION AY402518
 VERSION AY402518.1 GI:39758504
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 795)
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 Science 302 (5652), 1960-1963 (2003)
 JOURNAL
 PUBMED 14671302
 2 (bases 1 to 795)
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 JOURNAL
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source 1..795
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <1..>795
 /gene="CLDN18"
 /locus_tag="HCM1250"
 gene
 ORIGIN
 Query Match 28.6%; Score 607.6; DB 29; Length 795;
 Best Local Similarity 86.4%; Pred. No. 8e-116;
 Matches 686; Conservative 0; Mismatches 99; Indels 9; Gaps 1
 QY 90 ATGTCACACACATGCCAAGTGTGGCGTTCCTCTGTCTCATCTCTGGGGGTGCGCCGCGC 149
 Db 1 ATGGCCACACACATGCCAAGTGTGGCGTTCCTCTGTCTCATCTCTGGGGGTGCGCCGCGC 60
 QY 150 TGCATCGGGCCACCGGATGACATGTGGACGCCACCCAGGACCTGTACGACAAACCCCGTC 209
 Db 61 TGCATAGCGCCACCTGGGATGACATGTGGACCTCAAGA CTTGTATGACAAACCAAGTC 120
 QY 210 ACCTCGCTCTCCAGTACGAAGGGCTCTCGAGAGCTCGGTGAGGACAGAGTTTCAGGCTTC 269
 Db 121 ACCGCGCTGTTCCAGTATGAGGGCTCTCGAGAGATTGCTGACACAGAGCTCGGGTTC 180
 QY 270 ACCGAATGCAGGCCCTTAITTCACCATCTCTGGGACATTCACGCCATGCTGCAGGACGTGCA 329
 Db 181 ACCGAGTGCAGGCCCATCTTCCACCATCTCTGGGCTTCACGCCATGCTGCAGGACGTGCA 240
 QY 330 GGCCTGATGATCGTAGGATCGCTCGGGTGCCATTGGGCTCTCTGATTCATCTTTGCC 389
 Db 241 GCCTCATCATCGTGGGCAATGTTCTCGGGGTTCATCGGTATCTCTGCTTCATCTTCGCG 300

	QY	390	CTGAAATGCATCCGGATTGGCAGCATGGAGACTTGTGCCAAGACCAACATGACATCGACC	444
	Db	301	CTGAAGTGTCATTCGATTGGTAGCATGGATGACTCTGCCAAGGCCAAGATGACTCTGACT	360
	QY	450	TCCGGGATCATGTTTCATGTCTCAGGTCTTTGTGCAAATTGCTGGAGTGTCTCTGTGTTGCC	509
	Db	361	TCTGGGATCTTTGTTCACTCACTCCGGCATCTGTGCAATCATTTGGTGTGTCTGTGTTTGGC	420
	QY	510	AACATGCTGGTGACTTAACCTTCTGGATGTCCAAGCTAACATGTACA-----CCGGC	560
	Db	421	AACATGCTGGTGACCAACTTCTGGATGTCCAAGCTAACATGTACAAGCGGCATGGCGCGC	480
	QY	561	ATGGGTGGGATGFTGACAGCTGTTTCAGACCAAGGTACACATTTGGTGGCGCTCTGTTCGTG	620
	Db	481	ATGGGTGGCATGGTGTGACACCGTTTCAGACCAAGGTACACCTTTGGTGGAGCTCTGTTCGTG	540
	QY	621	GCTGGGTGCTGTGGAGGCTCACACTAAATTGGGGTGTGATGATGTGATCGCTTGCOCGG	680
	Db	541	GGCTGGGTGTCTGGAGGCTCACCCCTGATTTGGGGAGTGATGATGTGATCGCTTGCOCGT	600
	QY	681	GCCCTGGCACAGAAGAACCACTAACAGCCGTTTTCTTATCATGTCCTCAGGCCACACAGT	740
	Db	601	GCCCTGCACACCATGATGACAGCAACTTCAAGAGTGTCTTACCATGSCCTCTGGCCAAAAT	660
	QY	741	GTTCCTCAAGCTGGAGGCTTCAAGCGCAGCATCTGGCTTTGGGTTCACAAACCAAAAAAC	800
	Db	661	GTTCCTCAAGCTGGAGGCTTAAAGCGCCCTTAAAGCCAGCATCTGGCTTTGGGTTCACAAACCAAG	720
	QY	801	AAGAAATATACGATGGAGGTGCCCGACAGAGGACGAGGTACAATCTTATCTCTTCCAAAG	860
	Db	721	AAGAAATCTACGATGGGSGTGC CGCACAGAAAGCATGAACAGTCTCATCTTACCAAAG	780
	QY	861	CACGACTATGTGTA	874
	Db	781	TATGACTATGTGTA	794
<hr/>				
	RESULT 10			
	AY402517			
	LOCUS	Pan troglodytes CLDN18 gene, VIRTUAL TRANSCRIPT, partial sequence,	617 bp DNA linear GSS 15-DEC-2000	
	DEFINITION	genomic survey sequence.		
	ACCESSION	AY402517		
	VERSION	AY402517.1 GI:39758503		
	KEYWORDS	GSS.		
	SOURCE	Pan troglodytes (chimpanzee)		
	ORGANISM	Eukaryotes		
		Pan troglodytes		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.		
	REFERENCE	1 (bases 1 to 617)		
	AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G.J., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
	TITLE	Inferred nonneutral evolution from human-chimp-mouse orthologous gene trios		
	JOURNAL	Science 302 (5652), 1960-1963 (2003)		

14671302	2 (bases 1 to 617)
REFERENCE	Clark,A.G., Gnanouba,S., Nielson,R., Thomas,P., Kejaraiwal,A.,
AUTHORS	Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
	Perriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
	Adams,M.D. and Cargill,M.
TITLE	Direct Submission
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
	Rockville, MD 20850, USA
COMMENT	This sequence was made by sequencing genomic exons and ordering
	them based on alignment.
FEATURES	Location/Qualifiers
source	1. .617
	/organism="Pan troglodytes"
	/mol_type="genomic DNA"

gene /db_xref="taxon:9598"
<1..>617
/gene="CLDN18"
/locus_tag="HCM1250"

ORIGIN

Query Match 28.0%; Score 593.2; DB 29; Length 617;
Best Local Similarity 96.4%; Pred. No. 8.2e-113;
Matches 595; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 259 GTTCAGGCTTACCGAATGACGGCCCTATTTCCACCATCTCTGGAGCTTCAGCCATGCTGC 318
DB 1 GTTCAGGCTTACCGAATGACGGCCCTATTTCCACCATCTCTGGAGCTTCAGCCATGCTGC 60

QY 319 AGGAGTGGGAGCCCTGATGATGCTGAGGATCTCTGGAGCTTCAGCCATGCTGC 378
DB 61 AGGAGTGGGAGCCCTGATGATGCTGAGGATCTCTGGAGCTTCAGCCATGCTGC 120

QY 379 CCATCTTGGCCCTGAATGATGCTGAGGATCTCTGGAGCTTCAGCCATGCTGC 438
DB 121 CCATCTTGGCCCTGAATGATGCTGAGGATCTCTGGAGCTTCAGCCATGCTGC 180

QY 439 TGACACTGACCTCCGGAGTATGCTGATGCTGAGGATCTCTGGAGCTTCAGCCATGCTGC 498
DB 181 TGACACTGACCTCCGGAGTATGCTGATGCTGAGGATCTCTGGAGCTTCAGCCATGCTGC 240

QY 499 CTGTGTTGCCAATGATGCTGAGGATCTCTGGAGCTTCAGCCATGCTGC 558
DB 241 CTGTGTTGCCAATGATGCTGAGGATCTCTGGAGCTTCAGCCATGCTGC 300

QY 559 GCATGGTGGGATGCTGAGGATCTCTGGAGCTTCAGCCATGCTGC 618
DB 301 GCATGGTGGGATGCTGAGGATCTCTGGAGCTTCAGCCATGCTGC 360

QY 619 TGGGCTGGGATGCTGAGGATCTCTGGAGCTTCAGCCATGCTGC 678
DB 361 TGGGCTGGGATGCTGAGGATCTCTGGAGCTTCAGCCATGCTGC 420

QY 679 GGGGCTGGGATGCTGAGGATCTCTGGAGCTTCAGCCATGCTGC 738
DB 421 GGGGCTGGGATGCTGAGGATCTCTGGAGCTTCAGCCATGCTGC 480

QY 739 GTGTGCTTCAAGCTGGGATCTCTGGAGCTTCAGCCATGCTGC 798
DB 481 GTGTGCTTCAAGCTGGGATCTCTGGAGCTTCAGCCATGCTGC 540

QY 799 ACAAGAGATATACGATGAGGATCTCTGGAGCTTCAGCCATGCTGC 858
DB 541 ACAAGAGATATACGATGAGGATCTCTGGAGCTTCAGCCATGCTGC 600

QY 859 AGCAGCATGCTGTAA 875
DB 601 AGCAGCATGCTGTAA 617

RESULT 11

BQ016897
LOCUS BQ016897 627 bp mRNA linear EST 17-JUN-2002
DEFINITION UI-H-D10-aug-b-03-0-UI.s1 NCI CGAP_D10 Homo sapiens cDNA clone
IMAGE:5876258 3', mRNA sequence.
ACCESSION BQ016897
VERSION BQ016897.1 GI:19752174
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 627)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.

FEATURES
source

1..627

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5876258"

/tissue_type="Lung Focal Fibrosis"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NCI CGAP_D10"

/note="Organ: Lung; Vector: pT73-pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP_D10 is a cDNA library containing the following tissue(s): A pool of Lung Focal Fibrosis. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ATACGGGTC.

TAG TISSUE=lung with fibrosis

TAG LIB=UI-H-D10

TAG_SEQ=ATACGGGTC"

ORIGIN

Query Match 27.7%; Score 588.2; DB 12; Length 627;

Best Local Similarity 97.7%; Pred. No. 8.9e-112;

Matches 607; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 1102 TTTTCAATCTCTCAATTTCTTTTAAATATACTTTCTCTCTGATGAGAGATGCGT 1161

DB 7 TTTTATTTTAAATTTATTTTAAATATACTTTCTCTCTGATGAGAGATGCGT 66

QY 1162 TTTAATCTCTCTCAATTTTGAATTTAGACAGACTCCCTCTCTCTCTAGTCA 1221

DB 67 TTTAATCTCTCTCAATTTTGAATTTAGACAGACTCCCTCTCTCTAGTCA 126

QY 1222 TAAACCATTTGATGATCTTTCCAGCTATCCCAAGAAACTTTTGAAGGAAAGAG 1281

DB 127 TAAACCATTTGATGATCTTTCCAGCTATCCCAAGAAACTTTTGAAGGAAAGAG 186

QY 1282 TAGACCCAAAGATGTTATTTCTGCTTTTGAATTTTGTCTCCCAACCCCACTGCT 1341

DB 187 TAGACCCAAAGATGTTATTTCTGCTTTTGAATTTTGTCTCCCAACCCCACTGCT 246

QY 1342 AGTAATAACACTTACTTGAAGGAAAGCAATAAGAGAAAGATTTTGAATCTCTCCAGC 1401

DB 247 AGTAATAACACTTACTTGAAGGAAAGCAATAAGAGAAAGATTTTGAATCTCTCCAGC 306

QY 1402 CCATGATCTCGTCTTCTTACATGATCTTAAAGTTACCAACCAAGTCAATTTTCA 1461

DB 307 CCATGATCTCGTCTTCTTACATGATCTTAAAGTTACCAACCAAGTCAATTTTCA 366

QY 1462 GTTTGAGGCAACCAACCTTTCTACTGCTGTCATCTTTTATTTACAGCAACCACTT 1521

DB 367 GTTTGAGGCAACCAACCTTTCTACTGCTGTCATCTTTTATTTACAGCAACCACTT 426

Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
The following repetitive elements were found in this cDNA sequence: 1-39, >AT rich#Low complexity (matched complement) 500-562, >AT rich#Low complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

Location/Qualifiers

1..627

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5876258"

/tissue_type="Lung Focal Fibrosis"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NCI CGAP_D10"

/note="Organ: Lung; Vector: pT73-pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP_D10 is a cDNA library containing the following tissue(s): A pool of Lung Focal Fibrosis. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ATACGGGTC.

TAG TISSUE=lung with fibrosis

TAG LIB=UI-H-D10

TAG_SEQ=ATACGGGTC"

ORIGIN

Query Match 27.7%; Score 588.2; DB 12; Length 627;

Best Local Similarity 97.7%; Pred. No. 8.9e-112;

Matches 607; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 1102 TTTTCAATCTCTCAATTTCTTTTAAATATACTTTCTCTCTGATGAGAGATGCGT 1161

DB 7 TTTTATTTTAAATTTATTTTAAATATACTTTCTCTCTGATGAGAGATGCGT 66

QY 1162 TTTAATCTCTCTCAATTTTGAATTTAGACAGACTCCCTCTCTCTAGTCA 1221

DB 67 TTTAATCTCTCTCAATTTTGAATTTAGACAGACTCCCTCTCTCTAGTCA 126

QY 1222 TAAACCATTTGATGATCTTTCCAGCTATCCCAAGAAACTTTTGAAGGAAAGAG 1281

DB 127 TAAACCATTTGATGATCTTTCCAGCTATCCCAAGAAACTTTTGAAGGAAAGAG 186

QY 1282 TAGACCCAAAGATGTTATTTCTGCTTTTGAATTTTGTCTCCCAACCCCACTGCT 1341

DB 187 TAGACCCAAAGATGTTATTTCTGCTTTTGAATTTTGTCTCCCAACCCCACTGCT 246

QY 1342 AGTAATAACACTTACTTGAAGGAAAGCAATAAGAGAAAGATTTTGAATCTCTCCAGC 1401

DB 247 AGTAATAACACTTACTTGAAGGAAAGCAATAAGAGAAAGATTTTGAATCTCTCCAGC 306

QY 1402 CCATGATCTCGTCTTCTTACATGATCTTAAAGTTACCAACCAAGTCAATTTTCA 1461

DB 307 CCATGATCTCGTCTTCTTACATGATCTTAAAGTTACCAACCAAGTCAATTTTCA 366

QY 1462 GTTTGAGGCAACCAACCTTTCTACTGCTGTCATCTTTTATTTACAGCAACCACTT 1521

DB 367 GTTTGAGGCAACCAACCTTTCTACTGCTGTCATCTTTTATTTACAGCAACCACTT 426

```

QY 1522 CTAGGAGTTCTCTGAGCTCTCAGCTGAGTCTCTTTCTGTCGGGTGAGAAATTGTCC 1581
DB 427 CTAGGAGTTCTCTGAGCTCTCAGCTGAGTCTCTTTCTGTCGGGTGAGAAATTGTCC 486
QY 1582 CTAGATGAATGAGAAATTA-TTTTCTTTAAATTAAGTCTCTAAATAGTTAAATAAAT 1640
DB 487 CTAGATGAATGAGAAATTA-TTTTCTTTAAATTAAGTCTCTAAATAGTTAAATAAAT 546
QY 1641 AATGTTTAAATGATACACTATCTCTGGAATAGCTCACCCTACATGGGATA 1700
DB 547 AATGTTTAAATGATACACTATCTCTGGAATAGCTCACCCTACATGGGATA 606
QY 1701 GAAGGAATGAAAAATAATT 1721
DB 607 GAAGGAATGAAAAATAATT 627

RESULT 12
BI823718 743 bp mRNA linear EST 04-OCT-2001
LOCUS 603040838F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5181825 5',
DEFINITION mRNA sequence.
ACCESSION BI823718
VERSION BI823718.1 GI:15935268
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL 1 (bases 1 to 743)
COMMENT NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11453 row: k column: 10
High quality sequence stop: 554.
Location/Qualifiers
1. .743
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5181825"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/notes="Organ: pooled brain, lung, testis; Vector:
pCMV-SORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 Kb,
insert size range 1-3 Kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 26.0%; Score 552; DB 12; Length 743;
Best Local Similarity 95.4%; Pred. No. 3e-104;
Matches 666; Conservative 0; Mismatches 20; Indels 12; Gaps 9;

QY 28 CACACCTTCGGCAGCAGGCGGCGAGCTTCTCGCAGGCGGCGGCGGCGGCGGCGGCGG 87
DB 1 CACACCTTCGGCAGCAGGCGGCGGCGAGCTTCTCGCAGGCGGCGGCGGCGGCGGCGGCG 60
QY 88 TCATGTCCACCACCATGCACATGCCAAGTGTGGCGTTCTCTCTGTCATCTCTGGGCTGCGCG 147

```

```

DB 61 TCATGTACAAACACCATGCAGTGGTGGCGTCTCTCTGTCATCTCTGGGCTGCGCG 120
QY 148 GCTGATCGCGCCACCGGATGGACATGTGAGCACCAGGACCTGTAGGCAACCCCG 207
DB 121 GCTGATCGCGCCACCGGATGGACATGTGAGCACCAGGACCTGTAGGCAACCCCG 180
QY 208 TCACCTCCGTGTTCAGTACGAAAGGCTCTGAGAGAGCTGCTGAGGCGAGAGTTCAAGGCT 267
DB 181 TCACCTCCGTGTTCAGTACGAAAGGCTCTGAGAGAGCTGCTGAGGCGAGAGTTCAAGGCT 240
QY 268 TCACCGAATGACAGGCGCTTATTTTCCATCCTCGGACTTCCAGCCCATGCTGCGAGGCGAGTGC 327
DB 241 TCACCGAATGACAGGCGCTTATTTTCCATCCTCGGACTTCCAGCCCATGCTGCGAGGCGAGTGC 300
QY 328 GAGCCCTGATGATGTAGGCTGCTGGGTGCCA-TTGGGCTCTGCTGATCCATCTTT 386
DB 301 GAGCCCTGATGATGTAGGCTGCTGGGTGCCA-TTGGGCTCTGCTGATCCATCTTT 360
QY 387 GCCTGGAATGCATCCGCAATTCGAGCATGGAGGACTCTGCCAAGCCCAACATGAC-ACT 445
DB 361 GCCTGGAATGCATCCGCAATTCGAGCATGGAGGACTCTGCCAAGCCCAACATGAC-ACT 420
QY 446 GACCTCCGGGATCATGTTCATTTCTCAGGTCTTTGTGCAATTGTGTGGAGTGTCTGTGT 505
DB 421 GACCTCCGGGATCATGTTCATTTCTCAGGTCTTTGTGCAATTGTGTGGAGTGTCTGTGT 480
QY 506 TGCACCAATGCTGCTGACTAATCTTCTGGATTCACAGCTTAACATGACACCGGATGG 565
DB 481 TGCACCAATGCTGCTGACTAATCTTCTGGATTCACAGCTTAACATGACACCGGATGG 540
QY 566 T-GGGATGTGCAG-ACTGTTACAGCAGGTACACATTTGGTGGCG--TCTGTCTGTGG 621
DB 541 TGGGATGTGCAGAACTGTACAGCAGGTACACATTTAGTGGCCCATCTGTACGTGG 600
QY 622 GCTGGTCTGCTG-AGGCCCTCACATTAATTTGGGGTGTGATG-ATGTGATCG--CCTGC 677
DB 601 GCTGGTCTGCTGGAAGGACTCACACATTAATTTAGGTTGTGATGATGTCATCGGCTGCC 660
QY 678 CGGGGCTGGCAGCAGAA--GAACCACTACAAAGCC 713
DB 661 GGGGGCTTGACACCAAGAAAGAAACCACTACAAAGCC 698

RESULT 13
CB462343 681 bp mRNA linear EST 26-MAR-2003
LOCUS 722418 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION CB462343
ACCESSION CB462343
VERSION CB462343.1 GI:29268727
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 681)
Smith, R.P.L., Roberts, A.J., Ehternkamp, S.E., Chitko-McKown, C.G.,
Wray, J.E. and Keele, J.W.
A second set of bovine ESTs from pooled-tissue normalized libraries
Unpublished (2003)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: LAM8001 row: I column: 12
Seq primer: GTAATACGACTCACTATAGG.
Location/Qualifiers

```

FEATURES

```
source
1. .681
/organism="Bos taurus"
/mol_type="mrna"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARCK 6BOV"
/note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/endometrium."

ORIGIN
Query Match 24.8%; Score 525.8; DB 14; Length 681;
Best Local Similarity 86.9%; Pred. No. 8.5e-99;
Matches 578; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 44 GGAGGGCGGCGAGCTCTCGAGGGCGGAGGGCGGCGGCGGAGGATCATGTCCACCAAC 103
Db 17 GCGAGGCGAAGGCGGAGCTGGGCGGAGGAGCGAGAGCGGCGGCGGATGTCCACCAAC 76
QY 104 ATGCCAAGTGTGGCGTTCCCTCTCTCCATCTGGGGCTGGCGGCTGCATCGCGGCCAC 163
Db 77 ATGCCAAGTGTGGCGTTCCCTCTCTCCATCTGGGGCTGGCGGCTGCATCGTTGCCAC 136
QY 164 CGGATGACATGTGGAGACCCAGGACCTGTACGACAAACCCCGTCACTCGGTGTTCCA 223
Db 137 CGAGATGACATGTGGAGACCCAGGACCTATACGACAAACCCCGTCACTCGGTGTTCCA 196
QY 224 GTACGAGGCTCTGGAGAGCTGGTGAGGAGAGTTCAGGCTTCACGATGAGAGGCC 283
Db 197 GTATGAGGGCTCTGGCGAGTGGTGAGGAGAGTTCAGGCTTCACGATGAGAGGCC 256
QY 284 CTATTTCCACCTCTGGGACTTTCAGCCATCTGCAGGAGTGCAGGCGCTGATGATCGT 343
Db 257 CTACTTCACTCTGGGCTTGCAGGATCTGCAGGAGTGCAGGCGCTGATGATCGT 316
QY 344 AGGATCTCTGGGTGCCATTTGGCTCTCTGTATTCATCTTTTGGCCCTGAAATGATCCG 403
Db 317 GGGCATCTCTGATGTCTATTGGCTCTCTGTGGCCATCTTTGGCCCTGAAATGATCCG 376
QY 404 CATTGGCAGATGGAGGACTCTGCCAAGCCAAATGACATGACCTCCGGGATCATGTT 463
Db 377 TATGGCAACATGATGATGACTCCGCCAAGCCAAATGACATGACCTCCGGGATCATGTT 436
QY 464 CATTGTCTCAGGTCTTTGTGCAATTTCTGGAGTGTCTGTGTTTGCACATGCTGGTAC 523
Db 437 CATCATGCGAGTCTCTGTGCAATCTGCTGGAGTGTCTGTGTTTGCACATGCTGGTAC 496
QY 524 TAATCTTGGATGTCACAGTAACTATGACACCGGCGATGGGTGGGATGGTGCAGACTGT 583
Db 497 TAATCTTGGATGTCACAGTAACTATGACACCGGCGATGGGTGGGATGGTGCAGACTGT 556
QY 584 TCAGACAGGTACACATTTGTGGGCTCTGTCTGTGGGCTGGGTGGTGGAGGCTCAC 643
Db 557 TCAGACAGGTACACATTTGTGGGCTCTGTCTGTGGGCTGGGTGGTGGAGGCTCAC 616
QY 644 ACTAATTTGGGGTGTGATGTGATGTGATGCTGCGGCGGCGCTGGCACCAGAAAGAACAA 703
Db 617 GCTGATTTGGGGCGTGTGATGTGATGCTGCGGCGGCGCTGGGCGGCTGGGCGGAGAACAA 676
QY 704 CTACA 708
Db 677 CTACA 681

RESULT 14
BM837765
LOCUS
DEFINITION K-EST0113903 S9SNU601 Homo sapiens cDNA clone S9SNU601-77-B11 5',
mRNA sequence.
ACCESSION BM837765
VERSION BM837765.1 GI:19194174
```

```
KEYWORDS EST. Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (Bases 1 to 651)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 77 row: B column: 11
High quality sequence stop: 651.
Location/Qualifiers
1. .651
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="S9SNU601-77-B11"
/sex="M"
/tissue_type="Ascites"
/cell_type="Epithelial"
/lab_host="Top10F"
/clone_lib="S9SNU601"
/note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;
Site 2: XhoI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including SfiI
site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized with Superscript II using SfiI
oligo-dT primer. After first strand synthesis, RNA was
degraded by NaOH treatment and cDNA was amplified by PCR
reaction. The PCR products were digested with SfiI and
cloned into DraIII- digested pME18S-FL3 vector. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN
Query Match 24.5%; Score 519; DB 12; Length 651;
Best Local Similarity 100.0%; Pred. No. 2.2e-97;
Matches 519; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1270 GAAAGGAAGAGTAGACCAAGATGTTATTTCTGCTGTTGATTTGTTGTCCTCCACC 1329
Db 1 GAAAGGAAGAGTAGACCAAGATGTTATTTCTGCTGTTGATTTGTTGTCCTCCACC 60
QY 1330 CCCAACTTGGCTAGTAATAACACTTACTGAAGAAGAGCAATAAGAGAAAGATATTGT 1389
Db 61 CCCAACTTGGCTAGTAATAACACTTACTGAAGAAGAGCAATAAGAGAAAGATATTGT 120
QY 1390 AATCTCTCCAGCCCATGATCTCGGTTTCTTTTACTGACTGTGATCTTAAAGTTACCAACCA 1449
Db 121 AATCTCTCCAGCCCATGATCTCGGTTTCTTTTACTGACTGTGATCTTAAAGTTACCAACCA 180
QY 1450 AAGTCATTTTCAGTTTGAAGGCAACCAACCTTTTACTGCTGTGATCTTCTATTATAC 1509
Db 181 AAGTCATTTTCAGTTTGAAGGCAACCAACCTTTTACTGCTGTGATCTTCTATTATAC 240
QY 1510 AGCAACACCAATCTTAGAGATTTCCTGAGCTCTCCACTGGAGTCTCTTTTCTGTGCGGGT 1569
Db 241 AGCAACACCAATCTTAGAGATTTCCTGAGCTCTCCACTGGAGTCTCTTTTCTGTGCGGGT 300
```

```

QY 1570 CAGAAATTGTCCTAGATGAATGAGAAAATATATTTTTTAAATTTAAAGTCCTTAAATATAG 1629
DB 301 CAGAAATTGTCCTAGATGAATGAGAAAATATATTTTTTAAATTTAAAGTCCTTAAATATAG 360
QY 1630 TTTAAATAAATATATGTTTTAGTAAATGATACACTATCTCTGTGGAATAGCCTCACCCCT 1689
DB 361 TTTAAATAAATATATGTTTTAGTAAATGATACACTATCTCTGTGGAATAGCCTCACCCCT 420
QY 1690 ACATGTGGATAGAGGAAAATGAAAAATAATTCCTTTGACATTTGCTATATGTAATGTAATTTG 1749
DB 421 ACATGTGGATAGAGGAAAATGAAAAATAATTCCTTTGACATTTGCTATATGTAATGTAATTTG 480
QY 1750 TAAAGTCATGCTTAAAGTACAAATTCATGAAAAGCTCAC 1788
DB 481 TAAAGTCATGCTTAAAGTACAAATTCATGAAAAGCTCAC 519

RESULT 15
BM788945
LOCUS
DEFINITION
5'-mRNA sequence.
ACCESSION
BM788945
VERSION
BM788945.1 GI:19137177
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 653)
Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
Oh, K.J., Cheong, J.B., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
Kim, Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 2 row: E column: 11
High quality sequence stop: 653.
Location/Qualifiers
1. .653
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S19N665307-2-E11"
/sex="M"
/lab_host="Top10P'"
/clone_lib="S19N665307"
/notes="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoRI
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10P, by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

FEATURES
source

ORIGIN

Query Match 23.4%; Score 495.8; DB 12; Length 653;
Best Local Similarity 88.9%; Pred. No. 1.5e-92;

```

High quality sequence stop: 549.

FEATURES
source

1. 549
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S21SN520-20-C12"
/sex="P"
/tissue_type="Stomach"
/cell_type="Floating aggregates"
/lab_host="SNU-520"
/clone_lib="S21SN520"
/note="Organ: Stomach; Vector: pTZ18Rpl; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F, by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 20.5%; Score 434; DB 12; Length 549;
Best Local Similarity 100.0%; Pred. No. 1.1e-79;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1355 TACTGAAGAAGCAATAAGAGAAAGATATTGTGAATCTCTCCAGCCCATGATCTCGT 1414
Db 1 TACTGAAGAAGCAATAAGAGAAAGATATTGTGAATCTCTCCAGCCCATGATCTCGT 60

QY 1415 TTCTTACACGTGATCTTAAAGTTACCAACCAAGTCATTTTCAGTTTGAGGCAACC 1474
Db 61 TTCTTACACGTGATCTTAAAGTTACCAACCAAGTCATTTTCAGTTTGAGGCAACC 120

QY 1475 AAACCTTTCTACTGCTGTGACATCTCTTTATACAGCAACACCATCTTAGAGTTTCT 1534
Db 121 AAACCTTTCTACTGCTGTGACATCTCTTTATACAGCAACACCATCTTAGAGTTTCT 180

QY 1535 GAGCTCTCAGTGGAGTCTCTTCTGTCGGGGTCAAGAAATGTCCCTAGATGATGAG 1594
Db 181 GAGCTCTCAGTGGAGTCTCTTCTGTCGGGGTCAAGAAATGTCCCTAGATGATGAG 240

QY 1595 AAAATTATTTTAAATTAAGTCTAAATAGTTAAATATAATGTTTGTAGTAA 1654
Db 241 AAAATTATTTTAAATTAAGTCTAAATAGTTAAATATAATGTTTGTAGTAA 300

QY 1655 ATGATACATATCTCTGGAATAGCCTCACCCCTACATGATGGATAGAGAAATGAAAA 1714
Db 301 ATGATACATATCTCTGGAATAGCCTCACCCCTACATGATGGATAGAGAAATGAAAA 360

QY 1715 AATAATGCTTTGACATGCTATATGCTATGTTGTAAGTCATGCTTAAGTACAAATTC 1774
Db 361 AATAATGCTTTGACATGCTATATGCTATGTTGTAAGTCATGCTTAAGTACAAATTC 420

QY 1775 CATGAAGAAGCTCAC 1788
Db 421 CATGAAGAAGCTCAC 434

RESULT 17

BG538299 961 bp mRNA linear EST 03-APR-2001
LOCUS 602566769Fl NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4691417 5',
DEFINITION mRNA sequence.

ACCESSION

BG538299.1 GI:13530531

VERSION

EST.

KEYWORDS

Source Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 961)

AUTHORS

NIH-MGC http://mgc.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: CLONTECH Laboratories, Inc.

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LCM1510 row: i column: 18

High quality sequence stop: 578.

Location/Qualifiers

1..961

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4691417"

/lab_host="DH10B (T1 phage-resistant)"

/clone_lib="NIH_MGC_77"

/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1:

SfiI (ggccgctcgcc); Site 2: SfiI (ggccattatgccc); 5' and

3' adaptors were used in cloning as follows: 5' adaptor

sequence: 5'-CACGCCATTATGCCC-3' and 3' adaptor sequence:

5'-ATTCTAGAGCGGAGGGCGGACATG-dT(30)BN-3' (where B = A,

C, or G and N = A, C, G, or T). Average insert size 1.9

Kb (range 0.5-4.0 kb). 12/15 colonies contained inserts

by PCR. This library was enriched for full-length clones

and was constructed by Clontech Laboratories (Palo Alto,

CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 19.5%; Score 413.2; DB 12; Length 961;
Best Local Similarity 86.0%; Pred. No. 2.2e-75;
Matches 626; Conservative 0; Mismatches 73; Indels 29; Gaps 14;

QY 1028 GCCTCAGTCTCTGCTCTTAATATTCCACCAATAAAGAGCTGAGTATTATGAATTAGA 1087
Db 1 GCCTCAGTCTCTGCTCTTAATATTCCACCAATAAAGAGCTGAGTATTATGAATTAGA 60

QY 1088 GGCTATAGCTCACATTTTCAATCTCTATTTCTTTTAAATATAAATCTTCTACTCTGA 1147
Db 61 GGCTATAGCTCACATTTTCAATCTCTATTTCTTTTAAATATAAATCTTCTACTCTGA 120

QY 1148 TGAGAGATGTGGTTTT-----AATCTCTCTCACATTTTGTGATTTAGACAGATCC 1202
Db 121 TGAGAGTAATGTGGTTTTTCTACACTCTCTCTCACATTTTGTGATTTAGACAGATCC 180

QY 1203 CCCTCTTCTCTCTAGTCAATAAACCATTGATGATCTATTTCCAGCTTATCCCAAGAA 1262
Db 181 CCCTCTTCTCTCTAGTCAATAAACCATTGATGATCTATTTCCAGCTTATCCCAAGAA 240

QY 1263 AACTTTTTGAAA-GGAAAGAGTAGACCCCAAGATGT-ATTTTCTGCTGTT-TGAATTTTG 1319
Db 241 AACTTTTTGAAACGGAAGAGTAGACCCCAAGATGTTCATTTCTGCTGTTTCTGAATTTGT 300

QY 1320 TCTCCCAACCCCAACTTGGCTAGTAATA-AACACTTACTGAAGAGAGCAAT-AAAGAG 1377
Db 301 ACTCCCAAGCCCAACTTGGCTAGTAATAAACCACTTACTGAAGAGAGCAATCAACGAG 360

QY 1378 AAAGATATTTGTAATCTCTCCAGCCCATGATCTCGTGTCTTCTTACACTGTGATCTT-AA 1436
Db 361 AAAGATATTTGTAATCTCTCCAGCCCATGATCTCGTGTCTTCTTACACTGTGATCTTCAA 420

```

QY 1437 AGTTACCAACCAAGTCATT-TTCAGTTGAGGCAACCAAC--CTTTCTACTGCTGTT 1493
Db 421 AGTTACCAAGCAGAGTCATTCTTCAGTTTGAGGCAACCAACATTTTCTACTGCTGA 480
QY 1494 GACATCTCTTATTACAGCAACACCATCTAGGAGTTT-CCTGAGCTCTCCACTGGAGTC 1552
Db 481 GACATCTCTTATTACAGCAACACCATCTAGGAGTTTACCTGAGCTCTCCACTGGAGTC 540
QY 1553 CTCCTTTCTGT-CGCGGGTCAGA-AATTGTCCTAGAGTAATGAGAAATATTTTTTTTA 1610
Db 541 CTCCTTTCTGTCCACGGGTGAGCAATTTGTCCTAGAGTAATGAGAAATATTTATCCCAA 600
QY 1611 AATTAGTCCTAAATATAGTTAAATAAATAATGTTTGTAGTAAATGATACA-----1662
Db 601 TATTAGTCCTAAATATAGTTAAATAAATAATATCAATCTCTAACTAAACATGATACA 660
QY 1663 ---CTATCTCTGTGAATAGCTCAACCCCTACATGTGATAG-AAGGAATGAAAAATA 1718
Db 661 CTTATCTCTTGTGCATAGAGCTCAGCCCTACATGTGATAGACAGGAATGAAAAATC 720
QY 1719 ATTGCTTT 1726
Db 721 ATTGCTTT 728

RESULT 18
BX528145 567 bp mRNA linear EST 27-JUN-2003
LOCUS BX528145 Sugano mouse embryo mewa Mus musculus cDNA clone
DEFINITION IMAGP998P055838 ; IMAGE:2352628, mRNA sequence.
ACCESSION BX528145
VERSION BX528145.1 GI:32297508
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 567)
Heil,O., Ebert,L., Neubert,P., Peters,M., Radelof,U., Schneider,D.
and Korn,B.
Mouse UnigeneSet - RZPD2
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGP998P055838.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Mouse UnigeneSet - RZPD2 (RZPDLIB No.981)
http://www.rzpd.de/CloneCards/cgi-
bin/showLib.pl.cgi/response?libNo=981 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
sugf, Primer sequence: CTCTGCTCTAAAGCTGGC.
FEATURES
source
1..567
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGP998P055838 ; IMAGE:2352628"
/dev_stage="embryo, 14 dpc"
/lab_host="DH10B"
/clone_lib="Sugano mouse embryo mewa"
/notes="Vector: pME18S-FL3; Site 1: DraIII (CACTGTGTG);
Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed
with an oligo(dT) primer [ATGTGGCTTTTTTTTTTTTT];
double-stranded cDNA was ligated to a DraIII adaptor

```

```

[TTGTCGCTACTGG], digested and cloned into distinct DraIII
sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site
CACCATGTG). XhoI should be used to isolate the cDNA
insert. Size selection was performed to exclude fragments
<1.5kb. Library constructed by Dr. Sumio Sugano
(University of Tokyo Institute of Medical Science).
Custom primers for sequencing: 5' end primer
CTTCTGCTCTAAAGCTGGC and 3' end primer
CGACCTGTCAGCTCGAGCACA."
ORIGIN
Query Match 18.6%; Score 394; DB 13; Length 567;
Best Local Similarity 85.2%; Pred. No. 2.3e-71;
Matches 455; Conservative 0; Mismatches 70; Indels 9; Gaps 1;
QY 89 CATGTCCACCACCATGCGCAAGTGGTGGCGTTCCTCTGTCCATCTGGGGCTGGCGG 148
Db 34 CATGGCCACCCACGAGTGGTAGGGCTTCTCTGTCTCCCTCTGGGGTCTGGCGG 93
QY 149 CTGCATCGCGGCCACCGGATGCGACATGTGGAGCACCCAGGACCTGTACGACAAACCCGT 208
Db 94 CTGCATAGCCGCCACTGGGATGCGACATGTGGAGCACTCAAGACCTGTATGACAAACCA 153
QY 209 CACCTCCGTGTTCCAGTAGCAAGGGCTCTGGAGGAGCTGCTGGAGCGAGTTCAGGCTT 268
Db 154 CACCGCGTGTTCAGTAGTGAAGGGCTCTGGAGGAGTTCGCTGCAACAGAGCTCGGGTT 213
QY 269 CACCGAATGCGAGGCCCTATTTCACCATCTGGGACTTCCAGCCATGCTCCAGGCAGTGG 328
Db 214 CACCGAGTGCAGGCCCATATTCCACCATCTGGGGCTTCCAGGCATGCTCCAGGTATGC 273
QY 329 AGCCCTGATGATGCTGAGGATCGTCTGGGTGCCATTGGCCCTCTGGTATCCATCTTTC 388
Db 274 AGCCCTGATGATGCTGGGGCATTTGTTCTGGGGTCACTCGGTATCCTCTGTCATCTTC 333
QY 389 CTGAAATGCATCCGCAATGGCAGCATGGAGGACTCTGCAAGGCCAACATGACACTGAC 448
Db 334 CTGAAATGCATTCGCAATGGTAGCATGATGATGATGATGATGATGATGATGATGATG 393
QY 449 CTCGGGATCATGTTTCATGTTCTCAGGCTCTTGTGCAATGCTGGAGTGTCTGTGTTTC 508
Db 394 TTCTGGGATCTTGTTCATCATCTCCGGCATCTGTGCAATCAATTTGTTGTGTGTTTC 453
QY 509 CAACATGCTGTGACTTAACCTCTGGATGTCACAGCTAACATGTACA-----CCGG 559
Db 454 CAACATGCTGTGAGCAACCACTTCTGGATGTCACAGTAAACATGTACAGCGGATGGCG 513
QY 560 CATGGTGGGATGGTGCAGACTGTTTCAGACAGGTACACATTTGGTGGCGCTCT 613
Db 514 CATGGTGGCATGGTGCAGACCGCTTCAGACAGGTACACCTTTGGTGGCAGCTCT 567
RESULT 19
BX528145 627 bp mRNA linear EST 26-OCT-2001
LOCUS BX528145 Sugano mouse embryo mewa Mus musculus cDNA clone
DEFINITION IMAGP998P055838 ; IMAGE:2352628, mRNA sequence.
ACCESSION BX528145
VERSION BX528145.1 GI:16497436
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 627)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
Konno,H., Konda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
TITLE

```



```

QY 1856 GCCTGGGCAACATGAGAGCCCTCTCTTCAAAAATACAGAGAGAAAATACGCCAGT 1915
Db 363 GCCTGGGCAACATGAGAGCCCTCTCTTCAAAAATACAGAGAGAAAATACGCCAGT 422
QY 1916 CATGTGTGCATACACTGTAGTCCAGCATTCGGGAGGCTGAGTGGGAGATCACTTG 1975
Db 423 CATGTGTGCATACACTGTAGTCCAGCATTCGGGAGGCTGAGTGGGAGATCACTTG 482
QY 1976 AGCCAGGAGGCTTGGGGCTGAGTGAGCCATGATCACACCACTGCCTCCAGCCAGGTG 2035
Db 483 AGCCAGGAGGCTTGGGGCTGAGTGAGCCATGATCACACCACTGCCTCCAGCCAGGTG 542
QY 2036 ACATAGCGAGATCTGTCTTAAAAAATAAAAAATAAATAATGGAACACAGCAAGTCCT 2093
Db 543 ACATAGCGAGATCTGTCTTAAAAAATAAAAAATAAATAATGGAACACAGCAAGTCCT 600

RESULT 22
BU946638/6 70376950U1 RALUTXL01 Rattus norvegicus cdNA, mRNA sequence.
DEFINITION BU946638
ACCESSION BU946638
VERSION BU946638.1 GI:37701955
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 797)
Fu, G.K. and Stuve, L.L.
An improved method for the construction of cDNA libraries for
highly efficient DNA sequencing from the 3' end of expressed genes
Unpublished (2003)
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
FEATURES
    source
        1..797
            /organism="Rattus norvegicus"
            /mol_type="mRNA"
            /db_xref="taxon:10116"
            /tissue_type="Lung"
            /clone_lib="RALUTXL01"
            /note="Rat, Lung, Clofibrate, 250 mg/kg, 12hr-day 28
            sacrifice, SD, M/F, Pool"
ORIGIN
Query Match 17.7%; Score 374.6; DB 13; Length 797;
Best Local Similarity 80.4%; Pred. No. 2.4e-67;
Matches 530; Conservative 0; Mismatches 109; Indels 20; Gaps 7;

QY 427 CCAAGCCCAACATGACACTGACTCGGATCATGTTCAATGTTCTCAGTCTTTGGCAA 486
Db 736 CCAAGCCCAACATGACTC-GACTTCGGGGAATGTTCAATCTCA-GAGTCTGGCAA 679
QY 487 TTGCTGGAGTGTCTGTTTGGCAACATGCTGCTGACTAACTCTTGATGTCACAGCTA 546
Db 678 TCATTGGTGTCTGTG-TTGGCAACATGCTGTTACCACTTCTGGATGTCAC-GCAG 621
QY 547 ACATGTACACCGGATGGGTGGATGGTGACAGATGTTCCAGACCAAGGTACACATTTGGT 606
Db 620 ACATGTACACCGGATGGGTGGATGGTGACAGCGTTCCAGACCAAGGTACATTTGGT 561
QY 607 CGGCTGTGTTGGGCTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 666
Db 560 CAGCTGTGTTGGGCTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 501
QY 667 GCATGCGCTGCGGGGCGCTGGGACACAGAGAAACCAACTTCAAGCCGTGCTTATCATG 726
Db 500 GCATGCGTGGCGCTGACCTCTGATGATGACCGCAACTTCAAGCCGTGCTTATCATG 441

```

```

QY 727 CUTCAGGCCACAGTGTTCCTACAAGCCTGGAGGCTTCAAGCCACAGCACTGCTTTGGGT 786
Db 440 CTTCTGGCCAAATGTTGCTTACAAGCCTGGAGGCTTTAAGGCCACAGCACTGCTTTGGGT 381
QY 787 CCAACACCAAAAACAAGAAGATATACGATGGAGGTGCGCCGACAGAGCAGGTACAAAT 846
Db 380 CCAACACCAAAAACAAGAAGATATATGATGGGGGTGCGCCGACAGAGCAGGTACAAAT 321
QY 847 CTTATCCTTCCAAAGCAGCACTATGTGTAATGCTCTAAGACCTCTCAGC-----ACGGGC 900
Db 320 CTCATCCTACCAAGTACGACTATGTGTAGTGTCTTAAGACCGCCCCCAACATGTGTG 261
QY 901 GGAAGAAA-----CTCCCGAGAGCTCACCCAAAAACAAGAGATGCCATCTAGATTC 955
Db 260 AGAAGAAAACCTTCTCCCAAGGAAGCTCACCCCAAGCAAGGAGATCTTACCTTGTTC 201
QY 956 TTCTTGCTTTGACTCACAGCTGGAAGTTAGAAAAGCTCGATTTTCATCTTTGGAGGC 1015
Db 200 TTGTTGCTTTTGAAGTGTGAAAGTCTTAATTTTCATCTGTGGGGAGGC 141
QY 1016 CAAATGGTCTTAGCCTCAGTCTCTCTCTAAATATTCACCATAAACAGCTGAGTTA 1074
Db 140 TAGACAGTCTTGGCCACA-----CGTCTTAATATCTCCATCACAACACACGAGTTA 87

RESULT 23
BF091029 421 bp mRNA linear EST 19-OCT-2000
MR3-SN0036-110900-011-e08 SN0036 Homo sapiens cdNA, mRNA sequence.
DEFINITION BF091029
ACCESSION BF091029.1 GI:10896739
VERSION BF091029.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 421)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, P.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=6t2=MR3-SN0036-110
900-011-e08&t3=2000-09-11&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 421.
Location/Qualifiers
    1..421
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /dev_stage="Adult"
        /clone_lib="SN0036"
        /note="Organ: stomach normal; Vector: puc18; Site 1: SmaI;
        Site 2: SmaI; A mini-library was made by cloning products
        derived from ORESTES PCR (U.S. Letters Patent application

```

No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match 17.6%; Score 373.6; DB 10; Length 421;
Best Local Similarity 97.0%; Pred. No. 4.2e-67;
Matches 391; Conservative 0; Mismatches 10; Indels 2; Gaps 1;

QY 924 CCAAAACAAAGGAGATCCCATCAGATTTCTTCTGCTTTTGTATCAGCTGGAAT 983
DB 12 CACCCAAAACAGAGATCCCATCAGATTTCTTCT-CTTTGACTCAGCTGGAAT 69
QY 984 TAGAAAGCTCGATTTCTATCTTTGAGAGGCCAAATGCTTAGCTCAGTCTGTCT 1043
DB 70 TAGAAAGCTCGATTTCTATCTTTGAGAGGCCAAATGCTTAGCTCAGTCTGTCT 129
QY 1044 CTAAATATCCACCAATAAACAGCTCAGTTATTTATGATAGAGCTATAGCTACAT 1103
DB 130 CTAAATATCCACCAATAAACAGCTCAGTTATTTATGATAGAGCTATAGCTACAT 189
QY 1104 TTCAATCTCTATTTCTTTTAAATATAACTTTCTACTCTGATGAGAGAAATGTGTT 1163
DB 190 TTCAATCTCTATTTCTTTTAAATATAACTTTCTACTCTGATGAGAGAAATGTGTT 249
QY 1164 TAACTCTCTCACAATTTGATGATTTAGACAGACTCCCTCTCTCTCTAGTCAATA 1223
DB 250 TAACTCTCTCACAATTTGATGATTTAGACAGACTCCCTCTCTCTCTAGTCAATA 309
QY 1224 AACCCATTCATGATTTATTTCCAGCTATCCCAAGAAACCTTTTGAAGGAAGAGTA 1283
DB 310 AACCCATTCATGATTTATTTCCAGCTATTTCCCAAGAAACCTTTTGAAGGAAGAGTA 369
QY 1284 GACCCAAAGATGTTATTTCTGCTGTTTGAATTTCTCTCCCC 1326
DB 370 GACCCAAAGATGTTATTTCTGCTGTTTGAATTTCTCTCCCC 412

RESULT 24
BE841291/c
LOCUS MR0-SN0040-310500-007-all SN0040 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE841291
ACCESSION BE841291 GI:10273669
VERSION BE841291.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 379)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20203663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the PAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-MR0-SN0040-310

500-007-all&t3=2000-05-31&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 379.
Location/Qualifiers
1..379

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="SN0040"
/note="Organ: stomach normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match 17.0%; Score 360.8; DB 10; Length 379;
Best Local Similarity 99.5%; Pred. No. 1.9e-64;
Matches 362; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1211 CTCCTAGTCAATATAACCCATTGATGATCTATTTCCAGCTTATCCCAAGAAACTTTTG 1270
DB 367 CGCTAGTCAATATAACCCATTGATGATCTATTTCCAGCTTATCCCAAGAAACTTTTG 308
QY 1271 AAAGGAAGAGTAGACCCCAAGATGTTATTTCTGCTGTTTGAATTTCTCTCCCAACC 1330
DB 307 AAAGGAAGAGTAGACCCCAAGATGTTATTTCTGCTGTTTGAATTTCTCTCCCAACC 248
QY 1331 CCAACTTGTGCTAGTATAAACAATCTACTCAAGAAAGCAATAAGAGAAAGATATTGTA 1390
DB 247 CCAACTTGTGCTAGTATAAACAATCTACTCAAGAAAGCAATAAGAGAAAGATATTGTA 188
QY 1391 ATCTCTCCAGCCCATGATCTCGGTTTCTTACATCTGATCTTAAAGTTACCAACCAA 1450
DB 187 ATCTCTCCAGCCCATGATCTCGGTTTCTTACATCTGATCTTAAAGTTACCAACCAA 128
QY 1451 AGTCATTTCAGTTTGGAGCAACCAACCTTTCTAGCTGTTGATCTCTTATTACA 1510
DB 127 AGTCATTTCAGTTTGGAGCAACCAACCTTTCTAGCTGTTGATCTCTTATTACA 68
QY 1511 GCAACACCACTTCTAGGAGTTTCTGAGCTCTCCACCTGGAGTCTCTTTCTGTGCGGGTC 1570
DB 67 GCAACACCACTTCTAGGAGTTTCTGAGCTCTCCACCTGGAGTCTCTTTCTGTGCGGGTC 8
QY 1571 AGAA 1574
DB 7 AGAA 4

RESULT 25

N99013

LOCUS

N99013 420 bp mRNA linear EST 10-APR-1996
IMAGE:297598 5', mRNA sequence.

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 420)

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,

Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Matra, M.,

Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,

Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and

Wilson, R.

The WashU-Merck EST Project

Unpublished (1995)

CONTACT: Wilson RK

Washington University School of Medicine

444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LINL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.

Seq primer: mob.REGA+ET

High quality sequence stop: 227.

Location/Qualifiers

1..420

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="GDB:1242510"

/db_xref="taxon:9606"

/clone="IMAGE:297588"

/dev_stage="19 weeks"

/lab_host="DH10B (ampicillin resistant)"

/clone_lib="Soares_fetal_lung_NBH119W"

/note="Organ: lung; Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGGAGCGCGCAATTTTCTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M. Fatima Bonaldo. This library was constructed from the same fetus as the fetal heart library, Soares fetal heart NBH119W."

ORIGIN

Query Match 16.5%; Score 349.6; DB 14; Length 420;

Best Local Similarity 97.1%; Pred. No. 4.1e-62;

Matches 366; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 1192 AGACAGACTCCCTCTTCTCTAGTCAATAAACCCATTGATGATCTATTTCACGCTT 1251

Db 15 AGACAGACTCCCTCTTCTCTAGTCAATAAACCCATTGATGATCTATTTCACGCTT 74

QY 1252 ATCCCCAGAAACATTTTGAAGGAAAGAGTAGACCCCAAGATGTTATTTCTGCTGTTT 1311

Db 75 ATCCCCAGAAACATTTTGAAGGAAAGAGTAGACCCCAAGATGTTATTTCTGCTGTTT 134

QY 1312 GAATTTGTCTCCCAACCCCACTTGGCTAGTAATAAACACTTACTGAAGAAGCAAA 1371

Db 135 GAATTTGTCTCCCAACCCCACTTGGCTAGTAATAAACACTTACTGAAGAAGCAAA 194

QY 1372 TAAGAGAAAGATATTGTAATCTCTCGGCCCATGATCTCGGTTTCTTACACTGTGATC 1431

Db 195 TAAGAGAAAGATATTGTAATCTCTCGGCCCATGATCTCGGTTTCTTACACTGTGATC 254

QY 1432 TTAAGAGTTACCAACCAAGTCATTTTTCAGTTTGGGCAACCAACCTTTCTACTGCTG 1491

Db 255 TTAAGAGTTACCAACCAAGTCATTTTTCAGTTTGGGCAACCAACCTTTCTACTGCTG 314

QY 1492 TTGACATCTTCTTATTACAGCAACA -CAATTCTAGGAGTTTCTCTGAGCTCTCCACTGGAG 1550

Db 315 TTGACATCTTCTTATTACAGCAACAACCACTTCTAGNAGTTTCTCTGAGCTCTCCAGTGGAG 374

QY 1551 TCCTCTTTCTGTCGGG 1567

Db 375 TCCTCTTTCTGTCGGG 391

RESULT 26

BE841654

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

BE841654 458 bp mRNA linear EST 22-SEP-2000
MRI-SN0064-220500-005-d01 SN0064 Homo sapiens cDNA, mRNA sequence.
BE841654
BE841654.1 GI:10274032
EST.

SOURCE

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1. (bases 1 to 458)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=st2=MR1-SN0064-220

500-005-d01&t3=2000-05-22&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 260

High quality sequence stop: 338.

Location/Qualifiers

1..458

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="SN0064"

/note="Organ: stomach normal; Vector: puc18; Site 1: SmaI;

Site 2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the puc 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

ORIGIN

Query Match 16.4%; Score 346.8; DB 10; Length 458;

Best Local Similarity 88.1%; Pred. No. 1.6e-61;

Matches 400; Conservative 0; Mismatches 52; Indels 2; Gaps 2;

QY 1187 GATTTAGACAGACTCCCTCTTCTCTAGTCAATAAACCCATTGATGATCTATTTCCTG 1246

Db 3 GATTTTACAGACTAGCCCTCTTCTCTAGACAAATTAACCATGATGAATATTTCCTG 62

QY 1247 AGCTTATCCCCAGAAAACATTTTGAAGGAAAGTAGACCCCAAGATGTTATTTCTGCG 1306

Db 63 AGCTTCTCCCCAGAAAACGTTTGTAAAGCCCGCTGACCCACCGAGCAATTTCTGCG 122

QY 1307 TGTTTGAATTTGTCTCCCAACCCCACTTTGGCTAGTAATAACACTTACTTGAAGAAGA 1366

Db 123 TG-AGGAATGTGACTCCCAACCCCACTTTGGCTAGTAATAACACTTACTTGAAGAAGA 181

QY 1367 AGCAATAAGAGAAGATATTGTAATCTCTCCAGCCCATGATCTCGGTTTCTTACACTG 1426

Db 182 AGCCATCAGCGATAGTATTTCGAGATCTCTCCAGCCCATGATCTCGGTTTCTTACACTG 241

QY 1427 TGATCTTAAAGTTTACCAACCAAGTCATTTTTCAGTTTGAAGCAACCAACCTTTCTAC 1486

Db 242 GGAACCTTAAAGTTTACCAACCAAGTCATTTTTCAGTTTGAAGCAACCAACCTTTCTAC 301

QY 1487 TGCCTGTGACATCTTCTTATTACGCAACCACTTCTAGGAGTTTCTGAGCTTCTCCACT 1546

```

Db      302 TGCTGTGACATCTTCTATTACAGCAACACCACTTCTAGAGTTTCTCTGAGCTCTCCACT 361
QY      1547 GGAGTCCCTCTTCTGCGGGGTCAGAAATGTCCTAGATGATGAGAAATTA-TTTT 1605
Db      362 GGAGTCCCTCTTCTGCGGGGTTTCAGAAATGCGCTAGATGATGAGAAATTA-TTTT 421
QY      1606 TTTTAATTAAAGTCCTAAATATATAGTTAAATAAA 1639
Db      422 TTTTAATTAAAGGCCCAATATAGTTTCAATAAA 455

RESULT 27
LOCUS   BM836109
DEFINITION K-EST0111603 S9SNU601 Homo sapiens cDNA clone S9SNU601-65-G03 5', mRNA sequence.
ACCESSION BM836109
VERSION   BM836109.1 GI:19192518
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 341)
AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE     21C Frontier Korean EST Project 2001
JOURNAL   Unpublished (2002)
COMMENT   Contact: Kim YS
          Genome Research Center
          Korea Research Institute of Bioscience & Biotechnology
          52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
          Tel: +82-42-860-4470
          Fax: +82-42-860-4409
          Email: yongsung@mail.kribb.re.kr
          Plate: 65 row: G column: 03
          High quality sequence stop: 341.
          Location/Qualifiers
            1..341
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="S9SNU601-65-G03"
              /sex="M"
              /tissue_type="Ascites"
              /cell_type="Epithelial"
              /cell_line="SNU-601"
              /lab_host="Top10F"
              /clone_lib="S9SNU601"
              /note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI; Site 2: XhoI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including SfiI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized with Superscript II using SfiI oligo-dT primer. After first strand synthesis, RNA was degraded by NaOH treatment and cDNA was amplified by PCR reaction. The PCR products were digested with SfiI and cloned into DraIII- digested pME18S-FL3 vector. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN
Query Match 16.1%; Score 341; DB 12; Length 341;
Best Local Similarity 100.0%; Pred. No. 2.6e-60;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      874 AATGCTCTAAGACTCTCAGCAGCGGCGGAAGAACTCCCGGAGAGCTCACCCAAAAAC 933

```

```

Db      1 AATGCTCTAAGACTCTCAGCAGCGGCGGAAGAACTCCCGGAGAGCTCACCCAAAAAC 60
QY      934 AAGGAGATCCCATCTAGATTTCTTCTGCTTTTGTACTCAGCTGGAAGTTAGAAAGCC 993
Db      61 AAGGAGATCCCATCTAGATTTCTTCTGCTTTTGTACTCAGCTGGAAGTTAGAAAGCC 120
QY      994 TCGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCTCAGTCTCTGTCTCTAAATATTC 1053
Db      121 TCGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCTCAGTCTCTGTCTCTAAATATTC 180
QY      1054 CACCATAAAAACAGCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTCAATCCCTC 1113
Db      181 CACCATAAAAACAGCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTCAATCCCTC 240
QY      1114 TATTTCTTTTAAATATACTTTCTACTCTGATGAGAGGATGCTGTTTAACTCTCTCT 1173
Db      241 TATTTCTTTTAAATATACTTTCTACTCTGATGAGAGGATGCTGTTTAACTCTCTCT 300
QY      1174 CTCACATTTTGATGATTAGACAGACTCCCTCTCTCTCTCC 1214
Db      301 CTCACATTTTGATGATTAGACAGACTCCCTCTCTCTCC 341

RESULT 28
LOCUS   AW867300/c
DEFINITION MR3-SN0036-280400-009-f08 SN0036 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW867300
VERSION   AW867300.1 GI:8001352
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 412)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,D.F., de Souza,S.J. and Simpson,A.J.
TITLE     Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE  20202663
PUBMED   10737800
COMMENT   Contact: Simpson A.J.G.
          Laboratory of Cancer Genetics
          Ludwig Institute for Cancer Research
          Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
          Tel: +55-11-2704922
          Fax: +55-11-2707001
          Email: asimpson@ludwig.org.br
          This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
          (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=MR3-SN0036-280400-009-f08&t3=2000-04-28&t4=1)
          Seq primer: puc 18 forward
          High quality sequence start: 5
          High quality sequence stop: 412.
          Location/Qualifiers
            1..412
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /dev_stage="Adult"
              /clone_lib="SN0036"
              /note="Organ: stomach normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of

```

tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN	Query Match	16.0%;	Score 340.2;	DB 10;	Length 412;
	Best Local Similarity	94.6%;	Pred. No. 3.7e-60;		
	Matches 385;	Conservative 0;	Mismatches 18;	Indels 4;	Gaps 3;
Qy	912	CCGGAGAGCTCACCCCAAAAACAGAGGAGATCCCATCTAGATTTCTTCTTGCTTTTGACTC	971		
Db	407	CCGGGAGCTCACCCCAAAAACAGAGGAGATCCCATCTAGATTTCTTCTTGCTTTTGACTC	348		
Qy	972	ACAGCTGGAAAGTTAGAAAAGCCTCGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCCT	1031		
Db	347	ACAGCTGGAAAGTTAGAAAAGCCTCGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCCT	288		
Qy	1032	CAGTCT-CTGTCCTCTAAATATTCACCATAAACAGCTGAGTATTTATGAATTAGAGGC	1090		
Db	287	CAGTCTGCTGTCTCTTCATATTCACCATATTCAGCTGAGTATTTATGAATTAGAGGC	228		
Qy	1091	TATAGCTCACATTTTCAATCTCTATTTCTTTTTTAAATATAACTTTCTACTCTGATGA	1150		
Db	227	TATAGCTCACATTTTCTATCTCTATTTCTTTTTTATATAGAATTTCTACTCTGATGA	168		
Qy	1151	GAGAAATGTGTTTTAAATCTCTCTCTCAATTTTGTATTTAGACAGACTCCCGCTCTTC	1210		
Db	167	GAGAAATGTGTTTTAAATCTCTCTCTCAATTTTGTATTTAGACAGACTCCCGCTCTTC	108		
Qy	1211	CTCCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCAAGAAACTTTTG	1270		
Db	107	CTCCTAGTCAATAAACCCATTGATGATCTATTTCCC-GCTTATCCCAAGCAACTTTTGT	49		
Qy	1271	AAAGGAAGAGTAGACCCCAAGATGTTATTTCTGCTGTTTGAATTT	1317		
Db	48	AA--GGAAGAGTAGACCCCAAGATGTCATTTTCTGCTGTTTGAATTT	4		

RESULT 29	
EG937956	
LOCUS	402 bp mRNA linear EST 11-JUN-2001
DEFINITION	Bovine Abomasum cDNA Library Bos taurus cDNA 5', mRNA
ACCESSION	EG937956
VERSION	EG937956.1
KEYWORDS	GI:14337328
SOURCE	EST.
ORGANISM	Bos taurus (cow)
	Bos taurus
	Bos taurus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
	Bovidae; Bovinae; Bos.
REFERENCE	1. (bases 1 to 402)
AUTHORS	Moore,S.S., Hansen,C., Li,C., Fu,A., Meng,Y. and Li,G.
TITLE	cDNA's from bovine abomasum tissue
JOURNAL	Unpublished (2001)
COMMENT	Contact: Dr. Stephen Moore Beef Genomics Laboratory Dept of AFNS, University of Alberta 410 Agri/For, Dept of AFNS, U of A, Edmonton, AB, T6G 2P5, Canada Tel: 780 492 0169 Fax: 780 492 4265 Email: stephen.moore@ualberta.ca The sequence best matches gb:AF221069 (Homo sapiens Claudin-18 mRNA, complete cds) in main database at high score of 456.0 and E-value of 1e-126 PCR Primers FORWARD: M13 Forward BACKWARD: M13 Reverse Seq primer: T3 primer High quality sequence stop: 402 POLYA=No..
FEATURES	Location/Qualifiers
source	1. 402 /organism="Bos taurus"

```

/mol_type="mRNA"
/db_xref="taxon:9913"
/sex="Two males and one female mixed"
/tissue_type="Gastrointestinal tissue (GIT)"
/cell_type="Epithelial"
/dev_stage="Young adult"
/lab_host="XLI-BlueMRF+-strain"
/clone_lib="Bovine Abomasum cDNA Library"
/note="Organ: Abomasum; Vector: Uni-22APXR; Site_1: EcoR
I; Site_2: Xho I"

```

ORIGIN

Query Match	16.0%; Score 339.6; DB 12; Length 402;
Best Local Similarity	90.3%; Pred. No. 5e-60;
Matches 363; Conservative	0; Mismatches 39; Indels 0; Gaps 0;
Qy	385 TTGGCCCTGAAATGCAATCCGCAATGGCAGCATGGAGGACTCTGCCAAAGCCAAACATGACAC 444
Db	
1	TTGGCCCTGAATGCAATCCGATATGGCAACATGGATGACTCCGCCAAAGCCAAATGACAC 60
Qy	445 TGACCTCCGGGATCATAGTTCAATGTCACAGTCTTTGTGCAATGCTGGAGTGTCGTGTGT 504
Db	
61	TCACCTCCGGGATCATTTCAATCATCGAGGCTCTCTGTGCAATCGCTGGAGTGTCGTGTGT 120
Qy	505 TTGCCAAATGCTGGTGCACTAACTTCTCGAGATGCCACAGCTTAACATGTACACCGGCATGG 564
Db	
121	TTGCCAAATGCTGGTTACTAACTTCTGGATGTCACAGCCAGCATGTTTACCCAGCATGG 180
Qy	565 GTGGGATGGTGCAGACTGTGTCAGACCAAGGTACACATTTGGTGGCGGCTCTGTTCGTGGGCT 624
Db	
181	GGGGGATGGTGCAGACCGTTTCAGACCAAGGTACACCTTTGGTGGCGGCTCTGTTCGTGGGCT 240
Qy	625 GGGTGCCTGAGAGGCTCACACTAATTTGGGGGTGTGATGATGTGCATCGCCTGCCGGGGCC 684
Db	
241	GGGTGCCTGAGAGGCTCACCGTGAATGGGGGCGTGTGATGTGCATCGCCTGCCGGGGTC 300
Qy	685 TGGCACCAGAGAAGAAACCAACTACAAAGCCGGTTTCTTATCATGCTCAGGCCACAGTGTGG 744
Db	
301	TGGCCCCGAGGAAACCAACTACAAAGCCGTCTCTTACCATGCTCTAGGCCACACACGTGCG 360
Qy	745 CCTACAGCCTGGAGGCTTCAAGGCCACGACTGGCTTTGGGT 786
Db	
361	CCTACAGGCTGGAGGCTTCAAGGCCACGACGAGTGGCTTTGAGT 402

RESULT 30

CA436277/c	CA436277	493 bp	mRNA	linear	EST 08-NOV-2002
LOCUS	UI-H-DTO-aud-1-05-0-UI.s1	NCI CGAP DT0	Homo sapiens	CDNA clone	
DEFINITION	UI-H-DTO-aud-1-05-0-UI 3',	mRNA sequence.			
ACCESSION	CA436277				
VERSION	CA436277.1	GI:24800697			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 493)				
AUTHORS	NCI-CGAP	http://www.ncbi.nlm.nih.gov/ncicgap.			
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Robert Strausberg, Ph.D.				

Tissue Procurement: Dr. Jose Mercuriade
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-34, >AT rich#Low complexity (matched complement)
77-367, >Alu (matched complement)

Seq primer: M13 FORWARD		
POLYA=Yes.		
FEATURES	Location/Qualifiers	
source	1..493	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="UI-H-DT0-aud-1-05-0-UI"	
	/tissue_type="Metastatic Chondrosarcoma"	
	/dev_stage="Adult"	
	/lab_host="DH10B (Life Technologies)"	
	/clone_lib="NCI CGAP DT0"	
	/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP DT0 is a cDNA library containing the following tissue(s): Metastatic Chondrosarcoma in Lung. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AACTGTTTCGG."	
	TAG_TISSUE=lung metastatic chondrosarcoma	
	TAG_LIB=UI-H-DT0	
	TAG_SEQ=AACTGTTTCGG"	
ORIGIN		
Query Match	16.0%;	Score 339; DB 14; Length 493;
Best Local Similarity	100.0%;	Pred. No. 6.5e-60;
Matches 339;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY	1783	GCTCACCTGTAATCCTAGCACTTTGGAGGCTGAGGAGGAGGATCACTTGAGCCCGAG 1842
Db	355	GCTCACCTGTAATCCTAGCACTTTGGAGGCTGAGGAGGAGGATCACTTGAGCCCGAG 296
QY	1843	AAGTTGAGACTAGCTGGGCAACATGGAGAGCCCTGTCTCAAAAATACAGAGAGAA 1902
Db	295	AAGTTGAGACTAGCTGGGCAACATGGAGAGCCCTGTCTCAAAAATACAGAGAGAA 236
QY	1903	AAAATCAGCCAGTCATGGTGGGATACACCTGTAGTCCCGAGGAGGCTGAGGTG 1962
Db	235	AAAATCAGCCAGTCATGGTGGGATACACCTGTAGTCCCGAGGAGGCTGAGGTG 176
QY	1963	GGAGGATCACTTGAGCCCGAGGAGGTTGGGCTGCAGTGAGCCATGATCACACCACTGCA 2022
Db	175	GGAGGATCACTTGAGCCCGAGGAGGTTGGGCTGCAGTGAGCCATGATCACACCACTGCA 116
QY	2023	CTTCACGCCAGGTGACATAGCGAGATCCTGTCTTAAAAAATAAATAAATAAATGGAACA 2082
Db	115	CTTCACGCCAGGTGACATAGCGAGATCCTGTCTTAAAAAATAAATAAATAAATGGAACA 56
QY	2083	CAGCAAGTCTTAGGAAGTAGGTTAAACTTAATCTTTTAA 2121
Db	55	CAGCAAGTCTTAGGAAGTAGGTTAAACTTAATCTTTTAA 17
RESULT 31		
BM819720	552 bp	mRNA linear EST 06-MAR-2002
LOCUS	K-EST008789	S18N669761 Homo sapiens cDNA clone S18N669761-20-E04
DEFINITION	5', mRNA sequence.	
ACCESSION	BM819720	
VERSION	BM819720.1	GI:19176133
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 552)	

AUTHORS	Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.	
TITLE	21C Frontier Korean EST Project 2001	
JOURNAL	Unpublished (2002)	
COMMENT	Contact: Kim YS Genome Research Center Korea Research Institute of Bioscience & Biotechnology 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea Tel: +82-42-860-4470 Fax: +82-42-860-4409 Email: yongsaung@mail.kribb.re.kr Plate: 20 row: E column: 04 High quality sequence stop: 552. Location/Qualifiers 1..552 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="S18N669761-20-E04" /sex="F" /lab_host="Top10F" /clone_lib="S18N669761" /note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."	
ORIGIN		
Query Match	16.0%;	Score 339; DB 12; Length 552;
Best Local Similarity	100.0%;	Pred. No. 6.4e-60;
Matches 339;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY	1783	GCTCACACCTGTAATCCTAGCACATTTGGAGGCTGAGGAGGAGGATCACTTGAGCCCGAG 1842
Db	96	GCTCACACCTGTAATCCTAGCACATTTGGAGGCTGAGGAGGAGGATCACTTGAGCCCGAG 155
QY	1843	AAGTTGAGACTAGCTGGGCAACATGGAGAGCCCTGTCTCTACAAAATACAGAGAGAA 1902
Db	156	AAGTTGAGACTAGCTGGGCAACATGGAGAGCCCTGTCTCTACAAAATACAGAGAGAA 215
QY	1903	AAAATCAGCCAGTCATGGTGGGATACACCTGTAGTCCCGAGGATCCCGGAGGCTGAGGTG 1962
Db	216	AAAATCAGCCAGTCATGGTGGGATACACCTGTAGTCCCGAGGATCCCGGAGGCTGAGGTG 275
QY	1963	GGAGGATCACTTTGAGCCCGAGGAGGTTGGGGCTGCAGTGAGCCATGATCACACCACTGCA 2022
Db	276	GGAGGATCACTTTGAGCCCGAGGAGGTTGGGGCTGCAGTGAGCCATGATCACACCACTGCA 335
QY	2023	CTTCAGCCAGGTGACATAGCGAGATCCTGTCTTAAAAAATAAATAAATAAATGGAACA 2082
Db	336	CTTCAGCCAGGTGACATAGCGAGATCCTGTCTTAAAAAATAAATAAATAAATGGAACA 395
QY	2083	CAGCAAGTCTTAGGAAGTAGGTTAAACTTAATCTTTTAA 2121
Db	396	CAGCAAGTCTTAGGAAGTAGGTTAAACTTAATCTTTTAA 434
RESULT 32		
BM819706	567 bp	mRNA linear EST 06-MAR-2002
LOCUS	BM819706	

```

DEFINITION K-EST0087863 S18N669761 Homo sapiens cDNA clone S18N669761-20-C12
5', mRNA sequence.
ACCESSION BM819706
VERSION BM819706.1 GI:19176119
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 567)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 20 row: C column: 12
High quality sequence stop: 567.
FEATURES
Location/Qualifiers
source
1..567
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S18N669761-20-C12"
/sex="p"
/lab_host="Top10P"
/clone_lib="S18N669761"
/notes="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10P by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
ORIGIN
Query Match 16.0%; Score 339; DB 12; Length 567;
Best Local Similarity 100.0%; Pred. No. 6.4e-60;
Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1783 GCTCACACCTGTATCTTAGCATTCTGGAGCGCTGAGAGAGGATCACTTGAGCCGAG 1842
DB 111 GCTCACACCTGTATCTTAGCATTCTGGAGCGCTGAGAGAGGATCACTTGAGCCGAG 170
QY 1843 AAGTTCGAGACTAGCCTGGGCAACATGGAGAGCGCTGTCTCTCAAAATACAGAGAA 1902
DB 171 AAGTTCGAGACTAGCCTGGGCAACATGGAGAGCGCTGTCTCTCAAAATACAGAGAA 230
QY 1903 AAATATCGCAGTCACTGGTGCAATACCTGTAGTCCAGCATTCGGGAGCGCTGAGGTG 1962
DB 231 AAATATCGCAGTCACTGGTGCAATACCTGTAGTCCAGCATTCGGGAGCGCTGAGGTG 290
QY 1963 GGAGATCACTTGAGCCGAGGAGGTTGGGGCTCAGTGAGCCATGATCACACCACTGCA 2022
DB 291 GGAGATCACTTGAGCCGAGGAGGTTGGGGCTCAGTGAGCCATGATCACACCACTGCA 350
QY 2023 CTCAGCGAGGTGACATAGCGAGATCCTGTCTAAATAAATAAATAAATAAATAAATGGAACA 2082
|||||

```

```

DB 351 CTCACCCAGGTGACATAGCGAGATCCTGTCTAAATAAATAAATAAATAAATGGAACA 410
QY 2083 CAGCAAGTCTTAGGAGTAGGTTAAACCTAAATCTTTTAA 2121
|||||
DB 411 CAGCAAGTCTTAGGAGTAGGTTAAACCTAAATCTTTTAA 449
|||||

BM794650 570 bp mRNA linear EST 05-MAR-2002
K-EST0076037 S21SNU520 Homo sapiens cDNA clone S21SNU520-24-H09 5',
mRNA sequence.
ACCESSION BM794650
VERSION BM794650.1 GI:19142882
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 570)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 24 row: H column: 09
High quality sequence stop: 570.
FEATURES
Location/Qualifiers
source
1..570
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S21SNU520-24-H09"
/sex="p"
/tissue_type="Stomach"
/cell_type="Floating aggregates"
/lab_host="Top10P"
/clone_lib="S21SNU520"
/notes="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10P by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
ORIGIN
Query Match 16.0%; Score 338.8; DB 12; Length 570;
Best Local Similarity 81.4%; Pred. No. 7e-60;
Matches 464; Conservative 0; Mismatches 2; Indels 104; Gaps 2;

QY 1626 ATAGTTAAAAATAAATAGTTTTAGTAAAAATGATACACTATCTCTGTGAATAGCCTCAC 1685
|||||
DB 1 ATAGTTAAAAATAAATAGTTTTAGTAAAAATGATACACTATCTCTGTGAATAGCCTCAC 60
|||||

QY 1686 CCCTACATGTGGATAGAGGAATGAAAAATAATTCCTGTGACATGCTCTATATGGTAC 1745
|||||

```



```
Db 61 CCCTACATGTGGATAGAGCAATGAAATAATTCCTTTGACATTTGCTATATGCTAC 120
QY 1746 TTTTGTAAAGTCATGCTTAAGTACAAATTCATGAAGA----- 1782
Db 121 TTTTGTAAAGTCATGCTTAAGTACAAATTCATGAAGA----- 180
QY 1783 ----- 1782
Db 181 CTTTGAGGTCCTATGGCTCTGATGTGTACATGATAGTAAGTGTAAAGCCATGTAAGAAGTA 240
QY 1783 -----GCTCACACCTGTATCTAGCCTTTGGGAGGCTGAGGAG 1822
Db 241 AATAATGTCTGGGCACACAGTGGCTCAGCCCTGTAATCTAGCCTTTGGGAGGCTGAGGAG 300
QY 1823 GAAGGATCACTTGAAGCCCAAGATTGAGAGTAGCTGGGCAACATGAGAGCCCTGTC 1882
Db 301 GAAGGATCACTTGAAGCCCAAGATTGAGAGTAGCTGGGCAACATGAGAGCCCTGTC 360
QY 1883 TCTACAAAATACAGAGAGAGAAAATCAGCCAGTCATGCTGTCATACACCTGTATGCCAG 1942
Db 361 TCTACAAAATACAGAGAGAGAAAATCAGCCAGTCATGCTGTCATACCTGTATGCCAG 420
QY 1943 CATTCGGGAGGCTGA--GGTGGAGGATCACTTGGAGCCAGGAGGTTGGGGCTGTCAGTG 2001
Db 421 CATTCGGGAGGCTGAGGGTGGGAGGATCACTTGGAGCCAGGAGGTTGGGGCTGTCAGTG 480
QY 2002 AGCATGATCACACCATGCTCAGCTCCAGCCAGGAGTACATGCGAGATCTCTTAAAAAAA 2061
Db 481 AGCATGATCACACCATGCTCAGCTCCAGCCAGGAGTACATGCGAGATCTCTTAAAAAAA 540
QY 2062 TAAAAATAAATAATGGAACACAGCAAGTC 2091
Db 541 TAAAAATAAATAATGGAACACAGCAAGTC 570
```

```
RESULT 34
BX108271/c
LOCUS
DEFINITION
  BX108271 NCI CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2459693;
  IMAGE:2459693, mRNA sequence.
ACCESSION
  BX108271
VERSION
  BX108271.1 GI:27835103
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
  Radelof,U., Schneider,D. and Korn,B.
  Human Unigeneset - RZPD3
  Unpublished (2003)
  Contact: Ina Rolfs
  RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
  Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
  RZPD: IWAGp998M066117
  RZPDLIB: I.M.A.G.E. cDNA Clone Collection;
  Human Unigeneset - RZPD3 (RZPDLIB No.972)
  http://www.rzpd.de/CloneCards/cgi-
  bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
  RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
  Heubnerweg 6, D-14059 Berlin, Germany
  Tel: +49 30 32639 101
  Fax: +49 30 32639 111
  www.rzpd.de
  This clone is available royalty-free from RZPD;
  contact RZPD (clone@rzpd.de) for further information. Seq primer:
  M13u, Primer sequence: CCGTTGTAACGACGCCAGT.
FEATURES
  source
    1..458
    /organism="Homo sapiens"
    /mol_type="mRNA"
```

```
/db_xref="taxon:9606"
/clone="IMAGE:2459693"
/tissue_type="poorly differentiated adenocarcinoma with
  signet ring cell features"
/lab_host="DH10B"
/clone_lib="NCI CGAP Gas4"
/notes="Organ: stomach; Vector: pCMV-SPORT6; Site 1: SalI;
  Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
  Average insert size 1.69 kb. Life Technologies catalog #:
  11549-011"
ORIGIN
  Query Match 15.9%; Score 337.4; DB 13; Length 458;
  Best Local Similarity 99.7%; Pred. No. 1.4e-59;
  Matches 338; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
  QY 1783 GCTCACACCTGTATCTAGCCTTTGGGAGGCTGAGGAGGATCCTTGGAGCCAG 1842
  Db 351 GCTCACACCTGTATCTAGCCTTTGGGAGGCTGAGGAGGATCCTTGGAGCCAG 292
  QY 1843 AAGTTGAGACTAGCTGGGCAACATGAGAGCCCTGTCTTACAAAATACAGAGAA 1902
  Db 291 AAGTTGAGACTAGCTGGGCAACATGAGAGCCCTGTCTTACAAAATACAGAGAA 232
  QY 1903 AAAATCAGCCAGTCATGCTGGCATACACCTGTAGTCCAGCATTCGGGAGGCTGAGGTG 1962
  Db 231 AAAATCAGCCAGTCATGCTGGCATACACCTGTAGTCCAGCATTCGGGAGGCTGAGGTG 172
  QY 1963 GGAGGATCACTTTCAGCCAGGAGGTTGGGGCTGTCAGTGCATCATCACCACCTGCA 2022
  Db 171 GGAGGATCACTTTCAGCCAGGAGGTTGGGGCTGTCAGTGCATCATCACCACCTGCA 112
  QY 2023 CTCAGCCAGGTCACATAGCGAGATCTCTCTAAAAATAAATAAATAATATGAGACA 2082
  Db 111 CTCAGCCAGGTCACATAGCGAGATCTCTCTAAAAATAAATAAATAATATGAGACA 52
  QY 2083 CAGCAAGTCTAGCAAGTAGGTTAAATAAATAATCTTTAA 2121
  Db 51 CAGCAAGTCTAGCAAGTAGGTTAAATAAATAATCTTTAA 13
RESULT 35
BX108271/c
LOCUS
DEFINITION
  OP35B05.s1 Soares_NFL_T_OBC_S1 Homo sapiens cDNA clone
  IMAGE:1578801 3' similar to contains Alu repetitive element;; mRNA
  sequence.
ACCESSION
  AA961052
VERSION
  AA961052.1 GI:3127606
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 395)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  This clone is available royalty-free through LLNL; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Insert Length: 595 Std Error: 0.00
  Seq primer: -40m13 fwd. ET from Amersham
  High quality sequence stop: 372.
  Location/Qualifiers
    1..395
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:1578801"
    /lab_host="DH10B"
```

/clone lib="Soares NFL T GBC S1"
/notes="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NbHL19W, testis NHT, and B-cell
NCI CGAP GCBI) were mixed, and ss circles were made *in*
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 723095-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

Query Match	15.8%	Score 336	DB 9	Length 395
Best Local Similarity	100.0%	Prod. No. 2.8e-59		
Matches 336	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1783	GCTCACACCTGTAATCCTAGCACTTTGGGAGGCTTGAGGAGGAGGATCACTTGAGCCAG	1842	
Db	336	GCTCACACCTGTAATCCTAGCACTTTGGGAGGCTTGAGGAGGAGGATCACTTGAGCCAG	277	
Qy	1843	AAGTTCGAGACTAGCCTGGGCAACATCGAGAAAGCCCTGTCTCTCAAAAATACAGAGAA	1902	
Db	276	AAGTTCGAGACTAGCCTGGGCAACATCGAGAAAGCCCTGTCTCTCAAAAATACAGAGAA	217	
Qy	1903	AAATACGACGAGTCATGGTGGCATACACCTGTAGTCCAGCATTTCCGGAGGCTGAGGTG	1962	
Db	216	AAATACGACGAGTCATGGTGGCATACACCTGTAGTCCAGCATTTCCGGAGGCTGAGGTG	157	
Qy	1963	GGAGGATCACTTGAGCCAGGAGGTTGGGGCTGCAGTGAGCCATGATCACACCCTGCA	2022	
Db	156	GGAGGATCACTTGAGCCAGGAGGTTGGGGCTGCAGTGAGCCATGATCACACCCTGCA	97	
Qy	2023	CTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAAAATAAAAAATAATATGGAACA	2082	
Db	96	CTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAAAATAAAAAATAATATGGAACA	37	
Qy	2083	CAGCAAGTCTTAGGAGTAGGTTAAAAACTAATTCCT	2118	
Db	36	CAGCAAGTCTTAGGAGTAGGTTAAAAACTAATTCCT	1	

RESULT 36
 LOCUS BM785703
 DEFINITION K-EST0064196 S9SNU601 Homo sapiens CDNA clone S9SNU601-32-F11 5',
 585 bp mRNA linear EST 05-MAR-2002
 mRNA sequence.
 ACCESSION BM785703
 VERSION BM785703.1 GI:19133935
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 585)
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 TITLE 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished (2002)
 COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 32 row: E column: 11
 High quality sequence stop: 585.
 FEATURES
 location/Qualifiers
 1..585

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S9SNU601-32-E11"
/sex="M"
/tissue_type="Ascites"
/cell_type="Epithelial"
/cell_line="SNU-601"
/lab_host="Top10F"
/clone_lib="S9SNU601"
/note="Organ: Stomach; Vector: pME18-FL3; Site_1: XhoI; Site 2: XhoI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including Sfil site by treatment of 14 RNA ligase and the first strand cDNA was synthesized with Superscript II using Sfil oligo-dT primer. After first strand synthesis, RNA was degraded by NaOH treatment and cDNA was amplified by PCR reaction. The PCR products were digested with Sfil and cloned into pBr11- digested pME18-FL3 vector. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F, by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

```

Query Match	15.8%; Score 335.8; DB 12; Length 585;
Best Local Similarity	99.4%; Pred. No. 2.9e-59;
Matches 337; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
Qy 1783	GCTCACACCTGTAATCCTAGCACTTTGGGAGCGTGAGGAGGAGGATCACTTGAGCCCGAG 1842
Db	
Qy 137	GCTCAGCGCTGTAATCCTTAGCACTTTGGGAGCGTGAGGAGGAGGATCACTTGAGCCCGAG 196
Db	
Qy 1843	AACTTCGAGACTAGCCTCGGCAACATGGAGAAGCCCTGCTCTACAAAAATACAGAGAGAA 1902
Db	
Qy 197	AACTTCGAGACTAGCCTCGGCAACATGGAGAAGCCCTGCTCTACAAAAATACAGAGAGAA 256
Db	
Qy 1903	AAATCAGCCAGTCATGTGGGCATACACCTGTAGTCCCGAGCATTCGGGAGGCTGAGGTG 1962
Db	
Qy 257	AAATCAGCCAGTCATGTGGGCCTACACCTGTAGTCCCGAGCATTCGGGAGGCTGAGGTG 316
Db	
Qy 1963	GGAGGATCACTTGAGCCCGAGGAGGTTTGGGCTGCAGTCAGGCCATGATCACACCACTGCA 2022
Db	
Qy 317	GGAGGATCACTTGAGCCCGAGGAGGTTTGGGCTGCAGTCAGGCCATGATCACACCACTGCA 376
Db	
Qy 2023	CTCCAGCCAGGTGACATAGCGAGATCCTGCTCTAAAAAAAATAAAAAATAAATAATGAAACA 2082
Db	
Qy 377	CTCCAGCCAGGTGACATAGCGAGATCCTGCTCTAAAAAAAATAAAAAATAAATAATGAAACA 436
Db	
Qy 2083	CAGCAAGTCCTAGGAAGTAGGTTAAAACTAAATCTTTAA 2121
Db	
Qy 437	CAGCAAGTCCTAGGAAGTAGGTTAAACTAAATCTTTAA 475
Db	

RESULT 37	
BM791792	661 bp mRNA linear EST 05-MAR-2002
LOCUS	K-EST0071870 S21SNUS20 clone S21SNUS20-21-E06 5',
DEFINITION	mRNA sequence.
ACCESSION	BM791792.1 GI:19140024
VERSION	EST.
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 661)
AUTHORS	Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,O.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE	21C Frontier Korean EST Project 2001

FEATURES source

JOURNAL
COMMENT

Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 21 row: E column: 06
High quality sequence stop: 661.
Location/Qualifiers

FEATURES
source

1. .661
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S218NUS20-21-E06"
/sex="F"
/tissue_type="Stomach"
/cell_line="SNU-520"
/lab_host="Top10P"
/clone_lib="S218NUS20"
/notes="Organ: Stomach; Vector: pTZ19RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10P by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

Query Match 15.8%; Score 335.8; DB 12; Length 661;
Best Local Similarity 99.4%; Pred. No. 2.9e-59;
Matches 337; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1783 GCTCACCTGTATCTAGCACTTTGGAGGCTGGAGGAGGATCACTTGGCCCCAG 1842
Db 215 GCTCACCTGTATCTAGCACTTTGGAGGCTGGAGGAGGATCACTTGGCCCCAG 274
QY 1843 AAGTTCGAGACTAGCCCTGGGCAACATGGAGAAGCCCTGTCTTACAAAATACAGAGAA 1902
Db 275 AAGTTCGAGACTAGCCCTGGGCAACATGGAGAAGCCCTGTCTTACAAAATACAGAGAA 334
QY 1903 AAATCAGCCAGTCATGGTGGCATACACTGTGTAGTCCAGCATTCGGGGAGGTGAGGTG 1962
Db 335 AAATCAGCCAGTCATGGTGGCCCTACACCTGTGTAGTCCAGCATTCGGGGAGGTGAGGTG 394
QY 1963 GGGAGGATCACTTGGAGCCCGAGGAGTGGGGCTGCAGTGGCCATGATCACACCTGCA 2022
Db 395 GGGAGGATCACTTGGAGCCCGAGGAGTGGGGCTGCAGTGGCCATGATCACACCTGCA 454
QY 2023 CTCAGCCAGGTGACATAGCCAGATCTGTCTAAAAAATAAAAAATAAATATGGAACA 2082
Db 455 CTCAGCCAGGTGACATAGCCAGATCTGTCTAAAAAATAAAAAATAAATATGGAACA 514
QY 2083 CAGCAAGTCTTAGAAGTAGGTAAATACTAATCTTTAA 2121
Db 515 CAGCAAGTCTTAGAAGTAGGTAAATACTAATCTTTAA 553

RESULT 38
BB625263

LOCUS BB625263 633 bp mRNA linear EST 26-OCT-2001
DEFINITION BB625263 RIKEN full-length enriched, adult male cecum Mus musculus

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BB625263
BB625263.1 GI:16463563
EST.
Mus musculus (house mouse)
Mus musculus

REFERENCE
AUTHORS

Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 633)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)

TITLE
JOURNAL
COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
waji,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watanabe,M., Tameda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Kira,A.,
Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.,
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yananaka,I., and
Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.

FEATURES
source

Location/Qualifiers
1. .633
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="9130209K06"
/sex="male"
/tissue_type="cecum"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male cecum"
/notes="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGATCCAGAGCTCTTTTTTTTTTTTTTTT 3'] cDNA was
prepared by using trehalose thermo-activated reverse

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
 Contact: Robert Strauberg, Ph.D.
 Email: c9aps-remail.nih.gov
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 811 Std Error: 0.00
 Seq Primer: -40UP from Gibco
 High quality sequence stop: 428.

FEATURES
 source
 1. .483
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1740098"
 /dev_stage="19 weeks"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares fetal lung NBHL19W"
 /notes="Organ: lung; Vector: pT7T3D (Pharmacia) with a
 modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer
 [5'-TGTTACCAATGAGTGGGAGCGGCCAATTTTTTTTTTTT-3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT7T3 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M. Fatima Bonaldo. This library was constructed
 from the same fetus as the fetal heart library, Soares
 fetal heart NBHL19W."

ORIGIN
 Query Match 15.8%; Score 334.4; DB 9; Length 483;
 Best Local Similarity 99.7%; Pred. No. 5.9e-59;
 Matches 335; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1783 GCTCACACCTGTAATCTAGCACTTTGGGAGGCTGAGGAGGAGGATCACTTGAGCCCG 1842
 DB 336 GCTCACACCTGTAATCTAGCACTTTGGGAGGCTGAGGAGGAGGATCACTTGAGCCCG 277
 QY 1843 AAGTTCGAGCTAGCTGGGCAACATGAGAGCCCTGTCTCTACAAAATACAGAGAGAA 1902
 DB 276 AAGTTCGAGCTAGCTGGGCAACATGAGAGCCCTGTCTCTACAAAATACAGAGAGAA 217
 QY 1903 AAAATCAGCAGTCATGGTGGCATACACCTGTAGTCCAGCATTCGGGAGGCTGAGGTG 1962
 DB 216 AAAATCAGCAGTCATGGTGGCATACACCTGTAGTCCAGCATTCGGGAGGCTGAGGTG 157
 QY 1963 GGAGGATCACTTGAGCCCGAGGAGGTTGGGGCTGCAGTGGCCATGATCACACCTGCA 2022
 DB 156 GGAGGATCACTTGAGCCCGAGGAGGTTGGGGCTGCAGTGGCCATGATCACACCTGCA 97
 QY 2023 CTCACGCCAGGTGACATAGCAGATCCTGTCTAAAAAATAAAAAATAAATAATATGAACA 2082
 DB 96 CTCACGCCAGGTGACATAGCAGATCCTGTCTAAAAAATAAAAAATAAATAATATGAACA 37
 QY 2083 CAGCAAGTCTTAGAAGTAGTTAAACTAATCTTT 2118
 DB 36 CAGCAAGTCTTAGAAGTAGTTAAACTAATCTTT 1

RESULT 41
LOCUS BM829469
DEFINITION K-EST0102455 S9SNU601 Homo sapiens cDNA clone S9SNU601-60-D12 5',
 mRNA sequence.
ACCESSION BM829469
VERSION BM829469.1 GI:19185864
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 593)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.

TITLE
JOURNAL
COMMENT

21C Frontier Korean EST Project 2001
 Unpublished (2002)
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 60 row: D column: 12
 High quality sequence stop: 593.

FEATURES
 source

1. .593
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S9SNU601-60-D12"
 /sex="M"
 /tissue_type="Aescites"
 /cell_type="Epithelial"
 /cell_line="SNU-601"
 /lab_host="Top10F"
 /clone_lib="S9SNU601"
 /notes="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;
 Site 2: XhoI; The poly (A) + RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tabacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including Sfil
 site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized with Superscript II using Sfil
 oligo-dT primer. After first strand synthesis, RNA was
 degraded by NaOH treatment and cDNA was amplified by PCR
 reaction. The PCR products were digested with Sfil and
 cloned into DraIII- digested pME18S-FL3 vector. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."

ORIGIN

Query Match 15.8%; Score 334.2; DB 12; Length 593;
 Best Local Similarity 99.1%; Pred. No. 6.3e-59;
 Matches 336; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1783 GCTCACACCTGTAATCTAGCACTTTGGGAGGCTGAGGAGGAGGATCACTTGAGCCCG 1842
 DB 133 GCTCACACCTGTAATCTAGCACTTTGGGAGGCTGAGGAGGAGGATCACTTGAGCCCG 192
 QY 1843 AAGTTCGAGCTAGCTGGGCAACATGAGAGCCCTGTCTCTACAAAATACAGAGAGAA 1902
 DB 193 AAGTTCGAGCTAGCTGGGCAACATGAGAGCCCTGTCTCTACAAAATACAGAGAGAA 252
 QY 1903 AAAATCAGCAGTCATGGTGGCATACACCTGTAGTCCAGCATTCGGGAGGCTGAGGTG 1962
 DB 253 AAAATCAGCAGTCATGGTGGCATACACCTGTAGTCCAGCATTCGGGAGGCTGAGGTG 312
 QY 1963 GGAGGATCACTTGAGCCCGAGGAGGTTGGGGCTGCAGTGGCCATGATCACACCTGCA 2022
 DB 313 GGAGGATCACTTGAGCCCGAGGAGGTTGGGGCTGCAGTGGCCATGATCACACCTGCA 372
 QY 2023 CTCACGCCAGGTGACATAGCAGATCCTGTCTAAAAAATAAAAAATAAATAATATGAACA 2082
 DB 373 CTCACGCCAGGTGACATAGCAGATCCTGTCTAAAAAATAAAAAATAAATAATATGAACA 432
 QY 2083 CAGCAAGTCTTAGAAGTAGTTAAACTAATCTTTAA 2121
 DB 433 CAGCAAGTCTTAGAAGTAGTTAAACTAATCTTTAA 471

```

RESULT 42
AM027675/c
LOCUS
DEFINITION
AM027675
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 367)
NCI-CCAG http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAG clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco.
Location/Qualifiers
1..367
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2533931"
/tissue_type="poorly differentiated adenocarcinoma with
signed ring cell features"
/lab_host="DH10B"
/clone_lib="NCI CCAG Gas4"
/note="Organ: stomach; Vector: pCMV-SPORT6; Site 1: Sal.I;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.69 Kb. Life Technologies catalog #:
11549-011"

ORIGIN
Query Match
Best Local Similarity 15.7%; Score 333.8; DB 9; Length 367;
Matches 335; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1783 GTCACACCTGTAATCTAGCATTGTGGAGGCTGAGGAGGAGGATCACTTGGACCCAG 1842
DB GCTCAGCGCTGTATCTCTAGCATTTCGGAGGCTGAGGAGGAGGATCACTTGGACCCAG 278
QY 1843 AAGTTCGAGCTAGCTGGGCAACATGGAGAGCCCTGTCTCTACAAATACAGAGAA 1902
DB 277 AAGTTCGAGCTAGCTGGGCAACATGGAGAGCCCTGTCTCTACAAATACAGAGAA 218
QY 1903 AAATCAGCCAGTCATGTGGCATACCTGTAGTCCAGCATTCGGGAGGCTGAGGTG 1962
DB 217 AAATCAGCCAGTCATGTGGCATACCTGTAGTCCAGCATTCGGGAGGCTGAGGTG 158
QY 1963 GGAGGATCACTTGGACCCAGGAGGTTGGGGCTGAGGATGAGCCATGATCACACCACTGCA 2022
DB 157 GGAGGATCACTTGGACCCAGGAGGTTGGGGCTGAGGATGAGCCATGATCACACCACTGCA 98
QY 2023 CTCGAGCCAGGTGATACATGAGGATCCTGTCTAAAAAATAAAAAATAAATATGGAACA 2082
DB 97 CTCGAGCCAGGTGATACATGAGGATCCTGTCTAAAAAATAAAAAATAAATATGGAACA 38
QY 2083 CAGCAAGTCCTAGGAAGTAGGTAAACATAATCTTTT 2119
DB 37 CAGCAAGTCCTAGGAAGTAGGTAAACATAATCTTTT 1

```

```

RESULT 43
BB529462
LOCUS
DEFINITION
BB529462
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 649)
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,
Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J.,
Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,
Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)
On Jul 28, 2000 this sequence version replaced gi:9580920.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-3222
Fax: 81-45-503-3216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.
and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I.,
Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
Hayashizaki, Y.
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Location/Qualifiers
1..649
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="E030002G01"
/tissue_type="lung"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 0 day neonate"

```


Db 324 CCAGCCATGCTGCAAGCTGTACGAGCCCTGATGATGCTGGGCAATGTTCTGGGGTCATC 383
 Qy 366 GGCTCTCTGATTCATCTTTGCTTCCCTGAATGCATCCGATTCGACATCGAGACTCT 425
 Db 384 GGTATCCTCGTGTCCATCTTCGCTTGAAGTGCATTCGATGCTGGTAGCATGGATGACTCT 443
 Qy 426 GCCAAAGCAACATGACACTGACCTCCGGGATCATGTTCAITGCTCAGGTCCTTTGTGCA 485
 Db 444 GCCAAGGCCAAGATGACTCTGACTCTCTGGGATCTTGTTCATCACTCCGGCATCTGTGCA 503
 Qy 486 ATTGCTGAGTGTCTGCTGTTTCCCAACATGCTGGTGACTAACTTC 530
 Db 504 ATCATTGCTGTCTGCTGTTTGGCAACATGCTGGTGACCAACTTC 548

RESULT 45

CE210717 731 bp DNA linear GSS 25-SEP-2003
 LOCUS tigr-gss-dog-17000372715318 Dog Library Canis familiaris genomic,
 DEFINITION genomic survey sequence.

ACCESSION CE210717
 VERSION CE210717.1 GI:35366372

KEYWORDS GSS:

SOURCE Canis familiaris (dog)

ORGANISM Canis familiaris

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

1. (bases 1 to 731)

Kirkness, E.F., Baina, V., Halpern, A.L., Levy, S., Remington, K.,

Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and

Venter, J.C.

The dog genome: survey sequencing and comparative analysis

Science 301 (5641), 1898-1903 (2003)

22875432

14512627

Contact: Kirkness EF

The Institute for Genomic Research

Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,

Rockville, MD 20850, USA

Tel: 301-838-0200

Fax: 301-838-0208

Email: ekirknes@tigr.org

Class: shotgun.

Location/Qualifiers

source

1..731

/organism="Canis familiaris"

/mol_type="genomic DNA"

/strain="Standard Poodle"

/db_xref="taxon:9615"

/clone_lib="Dog Library"

/note="Site 1: BstXI; Libraries were prepared from

peripheral blood"

ORIGIN

Query Match 14.5%; Score 308; DB 29; Length 731;
 Best Local Similarity 72.2%; Pred. NO. 1.7e-53;
 Matches 536; Conservative 0; Mismatches 145; Indels 61; Gaps 8;

Qy 733 GCACAGTGTGCTACAGCCCTGGAGGCTTCAGGCCAGCACTGGCTTGGTCCAAACA 792

Db 20 GGCCCAATTCGCCTACAGCCAGGAGGCTTCAAGCCAGCACTGGCTTGGGCCCGACGA 79

Qy 793 CCNAAACAAAGAGATATACGATGAGGTGCCCGCAGAGGACGAGGTACATCTTATC 852

Db 80 GCNAAACAAAGAGATATACGATGAGGTGCCCGCAGAGGACGAGGTACATCTC 139

Qy 853 CTTCCAGACGACTATGTGTAAATGCTTAAGACCTCTCAGCACGGCGGGAAGAACTCC 912

Db 140 CTTCCAGTACGACTATGTGTAGCAATTCGAAGGCACTCAGCATGGCGGCAAGAACCC 199

Qy 913 C--GGAGAGCTACCCNAAACAAAGAGATCCCATCTAGATTTCTTCTGCTTTTGA 970

Db 200 CAATTAGAGCTCACTTGAAACAAAGGAATTTTACCTGACTCCTTATTTGCTTTTGA 259

Qy 971 CACAGCTGGAAAGTTAGAAAAGCCCTCGATTTTCATCTTTGGAGAGGCGCAAAATGGTCTTAGCC 1030
 Db 260 CACAGTTGAAAGTAAGAAAACCTTTGATTTTCATCGGAAGGAGGCGCAGATGCTCTGGGCC 319
 Qy 1031 TCAGTCTC--TGTCTCTAAATATTTCCACCATAAACAACAGCTGAGTTAT--TTATCAATTA 1085
 Db 320 ACTGTCTCTGTGTTTCTTAATGTTCCATCACAACAAACAGCTGAGTTATCAGATATGAATTA 379
 Qy 1086 GAGCTATAGCTCACATTTTCAA--TCCTCTATTCTTTTAAATATAAATTAACCTTTCTAC 1142
 Db 380 GATGCTACAGTTCAAGTTTCCATCTCTCGTGTGTTTTTTTTTAATTTAACCCTTTCTAA 439
 Qy 1143 TCTGATGAGAGAATGTGTTTTTAATCTCTCTCTCAATTTTGAATTTAGACAGACTCC 1202
 Db 440 TCTGATGATAAAATGTGTTTTGAATTCCTATCAGACAAAT----- 478
 Qy 1203 CCCTCTTCTCCTAGTCAATAAACCATTGATGATCTATTTCCTCCAGCTTATCCCAAGAA 1262
 Db 479 -----AGCTCATCGATGATCTATTTCCTCCAACTCATCCCCCAAGAA 517
 Qy 1263 AACTTTTCAAAGGAA-----AGAGTAGACCCAAAGATGTTTATTCTCTGCTGTTTGAATT 1316
 Db 518 CAATTCGAAAGTAAGAAAAGGAGTCCATCAAAAGACATCATTTCTTACTATTGATTT 577
 Qy 1317 TTGCTCTCCCAACCCCAACTTGGCTAGTAATAAACAACCTTACTGAAAGAGAACCAATAGA 1376
 Db 578 TAACTCTCCCA--GCCTACTTGGCTACTTAATAAACAACCTGATTTGAAGTAG-AGCTGATAGG 634
 Qy 1377 GAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTGTGATCTTAAA 1436
 Db 635 GAAAGATATTTGTAATGTTCTCCAGCTCATTAATCTGAGTTTTCTTACACTGGGATCTTGAC 694
 Qy 1437 AGTTACCAAAACCAAAAGTCAATTT 1458
 Db 695 ATTACCTGGACTGAAGCGATTT 716

Search completed: September 2, 2004, 17:25:50
 Job time : 3744 secs

;; PRIOR FILING DATE: 1998-10-14
;; PRIOR APPLICATION NUMBER: 60/104987
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105000
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105002
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105104
;; PRIOR FILING DATE: 1998-10-21
;; PRIOR APPLICATION NUMBER: 60/105169
;; PRIOR FILING DATE: 1998-10-22
;; PRIOR APPLICATION NUMBER: 60/105266
;; PRIOR FILING DATE: 1998-10-22
;; PRIOR APPLICATION NUMBER: 60/105693
;; PRIOR FILING DATE: 1998-10-26
;; PRIOR APPLICATION NUMBER: 60/105694
;; PRIOR FILING DATE: 1998-10-26
;; PRIOR APPLICATION NUMBER: 60/105807

Query Match 100.0%; Score 1357; DB 10; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTQVAFLLSILGLACIAATGMDMWSQDLYDNPTVSFQYEGLRSCVRSQSGF 60
Db 1 MSTTTQVAFLLSILGLACIAATGMDMWSQDLYDNPTVSFQYEGLRSCVRSQSGF 60
QY 61 TCRPFTTILGLPAMLQAVRALMIVGILGAILLSIFALKCIRIGSMEDSAKANWTLT 120
Db 61 TCRPFTTILGLPAMLQAVRALMIVGILGAILLSIFALKCIRIGSMEDSAKANWTLT 120
QY 121 SGIMFTVSGLCATAGSVFANMLVTNFMSTANNYTGMGWQVQVQTRTYTFFGAALFVGWV 180
Db 121 SGIMFTVSGLCATAGSVFANMLVTNFMSTANNYTGMGWQVQVQTRTYTFFGAALFVGWV 180
QY 181 AGGLTLIGVVMCIACRGLAPEETNYKAVSYHASGSHVAYKPGCFKASTFGSNTKNKKI 240
Db 181 AGGLTLIGVVMCIACRGLAPEETNYKAVSYHASGSHVAYKPGCFKASTFGSNTKNKKI 240
QY 241 YDGGARTEDEVQSPSKHDYV 261
Db 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 2
US-10-206-915-356
; Sequence 356, Application US/10206915
; Publication No. US20040029221A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C513
; CURRENT APPLICATION NUMBER: US/10/206,915
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17

;; PRIOR APPLICATION NUMBER: 60/063120
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/063121
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/063486
;; PRIOR FILING DATE: 1997-10-21
;; PRIOR APPLICATION NUMBER: 60/063540
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/063541
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/063544
;; PRIOR FILING DATE: 1997-10-28
;; Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 612
;; SEQ ID NO 356
;; LENGTH: 261
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-206-915-356

Query Match 100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTQVAFLLSILGLACIAATGMDMWSQDLYDNPTVSFQYEGLRSCVRSQSGF 60
Db 1 MSTTTQVAFLLSILGLACIAATGMDMWSQDLYDNPTVSFQYEGLRSCVRSQSGF 60
QY 61 TCRPFTTILGLPAMLQAVRALMIVGILGAILLSIFALKCIRIGSMEDSAKANWTLT 120
Db 61 TCRPFTTILGLPAMLQAVRALMIVGILGAILLSIFALKCIRIGSMEDSAKANWTLT 120
QY 121 SGIMFTVSGLCATAGSVFANMLVTNFMSTANNYTGMGWQVQVQTRTYTFFGAALFVGWV 180
Db 121 SGIMFTVSGLCATAGSVFANMLVTNFMSTANNYTGMGWQVQVQTRTYTFFGAALFVGWV 180
QY 181 AGGLTLIGVVMCIACRGLAPEETNYKAVSYHASGSHVAYKPGCFKASTFGSNTKNKKI 240
Db 181 AGGLTLIGVVMCIACRGLAPEETNYKAVSYHASGSHVAYKPGCFKASTFGSNTKNKKI 240
QY 241 YDGGARTEDEVQSPSKHDYV 261
Db 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 3
US-10-199-670-356
; Sequence 356, Application US/10199670
; Publication No. US20040033560A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C401
; CURRENT APPLICATION NUMBER: US/10/199,670
; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250

;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/063120
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/063121
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/063121
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/063486
;; PRIOR FILING DATE: 1997-10-21
;; PRIOR APPLICATION NUMBER: 60/063540
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/063541
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/063544
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 612
;; SEQ ID NO 356
;; LENGTH: 261
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-201-858-356

Query Match 100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCCQVAFLLSILGLAGCIAATGMDMNSTQDLYDNPTSVFQYEGLMRSCVRQSSGF 60
DB 1 MSTTTCCQVAFLLSILGLAGCIAATGMDMNSTQDLYDNPTSVFQYEGLMRSCVRQSSGF 60
QY 61 TECRPYFTILGLPAMQAVRALMIVIGVLGAIGLLVSTFALKCIRIGSMEDSAKANMTLT 120
DB 61 TECRPYFTILGLPAMQAVRALMIVIGVLGAIGLLVSTFALKCIRIGSMEDSAKANMTLT 120
QY 121 SGIMFIVSGLCAIAGSVFANMLVTNFWNSTANMTYTGGMVQVOTVTRYTTCGAALFVGWV 180
DB 121 SGIMFIVSGLCAIAGSVFANMLVTNFWNSTANMTYTGGMVQVOTVTRYTTCGAALFVGWV 180
QY 181 AGGLTLIGVVMCIACRGLAPEETNYKAVSHASGHVSVAIKPGGFKASTGFGSNTKNKI 240
DB 181 AGGLTLIGVVMCIACRGLAPEETNYKAVSHASGHVSVAIKPGGFKASTGFGSNTKNKI 240
QY 241 YDGGARTEDEVQSPSKHDYV 261
DB 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 5

US-10-205-890-356
; Sequence 356, Application US/10205890
; Publication No. US20040048334A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C519
; CURRENT APPLICATION NUMBER: US/10/205,890
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266

;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/063120
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/063121
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/063486
;; PRIOR FILING DATE: 1997-10-21
;; PRIOR APPLICATION NUMBER: 60/063540
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/063541
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/063544
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 612
;; SEQ ID NO 356
;; LENGTH: 261
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-199-670-356

Query Match 100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCCQVAFLLSILGLAGCIAATGMDMNSTQDLYDNPTSVFQYEGLMRSCVRQSSGF 60
DB 1 MSTTTCCQVAFLLSILGLAGCIAATGMDMNSTQDLYDNPTSVFQYEGLMRSCVRQSSGF 60
QY 61 TECRPYFTILGLPAMQAVRALMIVIGVLGAIGLLVSTFALKCIRIGSMEDSAKANMTLT 120
DB 61 TECRPYFTILGLPAMQAVRALMIVIGVLGAIGLLVSTFALKCIRIGSMEDSAKANMTLT 120
QY 121 SGIMFIVSGLCAIAGSVFANMLVTNFWNSTANMTYTGGMVQVOTVTRYTTCGAALFVGWV 180
DB 121 SGIMFIVSGLCAIAGSVFANMLVTNFWNSTANMTYTGGMVQVOTVTRYTTCGAALFVGWV 180
QY 181 AGGLTLIGVVMCIACRGLAPEETNYKAVSHASGHVSVAIKPGGFKASTGFGSNTKNKI 240
DB 181 AGGLTLIGVVMCIACRGLAPEETNYKAVSHASGHVSVAIKPGGFKASTGFGSNTKNKI 240
QY 241 YDGGARTEDEVQSPSKHDYV 261
DB 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 4

US-10-201-858-356
; Sequence 356, Application US/10201858
; Publication No. US20040038337A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C464
; CURRENT APPLICATION NUMBER: US/10/201,858
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18

; PRIOR FILING DATE: 1997-09-18
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/063120
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/063121
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/063486
 ; PRIOR FILING DATE: 1997-10-21
 ; PRIOR APPLICATION NUMBER: 60/063540
 ; PRIOR FILING DATE: 1997-10-28
 ; PRIOR APPLICATION NUMBER: 60/063541
 ; PRIOR FILING DATE: 1997-10-28
 ; PRIOR APPLICATION NUMBER: 60/063544
 ; PRIOR FILING DATE: 1997-10-28
 ; Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 356
 ; LENGTH: 261
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-205-890-356

Query Match 100.0%; Score 1357; DB 12; Length 261;
 Best Local Similarity 100.0%; Pred. No. 2.7e-132;
 Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSTTTCCQVAFILSLGLAGCIAATGMDMSTQDLYDNPTVSFQYEGLRWSCVROSSGF 60
 Db 1 MSTTTCCQVAFILSLGLAGCIAATGMDMSTQDLYDNPTVSFQYEGLRWSCVROSSGF 60
 QY 61 TCRPYFTILGLPAMLOAVRALMIVGILGAILLVSIFALKCIRIGSMEDSAKANMTLT 120
 Db 61 TCRPYFTILGLPAMLOAVRALMIVGILGAILLVSIFALKCIRIGSMEDSAKANMTLT 120
 QY 121 SGIMFTVSGLCATAGSVFANMLVTNFWMTANNMTGMGMVQTVQTRYTFGAALFVGWV 180
 Db 121 SGIMFTVSGLCATAGSVFANMLVTNFWMTANNMTGMGMVQTVQTRYTFGAALFVGWV 180
 QY 181 AGGLTLIGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
 Db 181 AGGLTLIGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
 QY 241 YDGGARTEDEVQSPSKHDYV 261
 Db 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 6

US-10-208-024-356
 ; Sequence 356, Application US/10208024
 ; Publication No. US20040048335A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3430R1C538
 ; CURRENT APPLICATION NUMBER: US/10/208,024
 ; CURRENT FILING DATE: 2002-07-29
 ; PRIOR APPLICATION NUMBER: 10/052586
 ; PRIOR FILING DATE: 2002-01-15
 ; PRIOR APPLICATION NUMBER: 60/059263
 ; PRIOR FILING DATE: 1997-09-18

; PRIOR APPLICATION NUMBER: 60/059266
 ; PRIOR FILING DATE: 1997-09-18
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/063120
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/063121
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/063486
 ; PRIOR FILING DATE: 1997-10-21
 ; PRIOR APPLICATION NUMBER: 60/063540
 ; PRIOR FILING DATE: 1997-10-28
 ; PRIOR APPLICATION NUMBER: 60/063541
 ; PRIOR FILING DATE: 1997-10-28
 ; PRIOR APPLICATION NUMBER: 60/063544
 ; PRIOR FILING DATE: 1997-10-28
 ; Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 356
 ; LENGTH: 261
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-208-024-356

Query Match 100.0%; Score 1357; DB 12; Length 261;
 Best Local Similarity 100.0%; Pred. No. 2.7e-132;
 Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSTTTCCQVAFILSLGLAGCIAATGMDMSTQDLYDNPTVSFQYEGLRWSCVROSSGF 60
 Db 1 MSTTTCCQVAFILSLGLAGCIAATGMDMSTQDLYDNPTVSFQYEGLRWSCVROSSGF 60
 QY 61 TCRPYFTILGLPAMLOAVRALMIVGILGAILLVSIFALKCIRIGSMEDSAKANMTLT 120
 Db 61 TCRPYFTILGLPAMLOAVRALMIVGILGAILLVSIFALKCIRIGSMEDSAKANMTLT 120
 QY 121 SGIMFTVSGLCATAGSVFANMLVTNFWMTANNMTGMGMVQTVQTRYTFGAALFVGWV 180
 Db 121 SGIMFTVSGLCATAGSVFANMLVTNFWMTANNMTGMGMVQTVQTRYTFGAALFVGWV 180
 QY 181 AGGLTLIGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
 Db 181 AGGLTLIGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKI 240
 QY 241 YDGGARTEDEVQSPSKHDYV 261
 Db 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 7

US-10-201-853-356
 ; Sequence 356, Application US/10201853
 ; Publication No. US20040053358A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3430R1C465
 ; CURRENT APPLICATION NUMBER: US/10/201,853
 ; CURRENT FILING DATE: 2002-07-23
 ; PRIOR APPLICATION NUMBER: 10/052586
 ; PRIOR FILING DATE: 2002-01-15
 ; PRIOR APPLICATION NUMBER: 60/059263

```
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-745-118

Query Match      100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTITTCQVAVFLLSILGLAGCIAATGMDMWSMTQDLYDNPNVTSVFQYEGLMRSCVRQSSGF 60
Db 1 MSTITTCQVAVFLLSILGLAGCIAATGMDMWSMTQDLYDNPNVTSVFQYEGLMRSCVRQSSGF 60
QY 61 TECRPYFTILGLPAMLOAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TECRPYFTILGLPAMLOAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
QY 121 SGIMFIVSGLCIAAGSVFANMLVTNFMWSTANMYTGMGMVQVQVQVQVQVQVQVQVQV 180
Db 121 SGIMFIVSGLCIAAGSVFANMLVTNFMWSTANMYTGMGMVQVQVQVQVQVQVQVQV 180
QY 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHVSVAIKPGGFKASTGFGSNTKNKKI 240
Db 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHVSVAIKPGGFKASTGFGSNTKNKKI 240
QY 241 YDGGARTEDEVQSPSKHDYV 261
Db 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 9
US-09-940-227-76
; Sequence 76, Application US/09940227
; Publication No. US20030017468A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Sei Yu
; APPLICANT: Macina, Roberto
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Hervé
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific
; FILE OF INVENTION: Genes
; FILE REFERENCE: DEX-0230
; CURRENT APPLICATION NUMBER: US/09/940,227
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,378
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-940-227-76

Query Match      100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTITTCQVAVFLLSILGLAGCIAATGMDMWSMTQDLYDNPNVTSVFQYEGLMRSCVRQSSGF 60
Db 1 MSTITTCQVAVFLLSILGLAGCIAATGMDMWSMTQDLYDNPNVTSVFQYEGLMRSCVRQSSGF 60
QY 61 TECRPYFTILGLPAMLOAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TECRPYFTILGLPAMLOAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
QY 121 SGIMFIVSGLCIAAGSVFANMLVTNFMWSTANMYTGMGMVQVQVQVQVQVQVQVQV 180
Db 121 SGIMFIVSGLCIAAGSVFANMLVTNFMWSTANMYTGMGMVQVQVQVQVQVQVQVQV 180
QY 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHVSVAIKPGGFKASTGFGSNTKNKKI 240
Db 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHVSVAIKPGGFKASTGFGSNTKNKKI 240
```

```
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 356
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-201-853-356

Query Match      100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTITTCQVAVFLLSILGLAGCIAATGMDMWSMTQDLYDNPNVTSVFQYEGLMRSCVRQSSGF 60
Db 1 MSTITTCQVAVFLLSILGLAGCIAATGMDMWSMTQDLYDNPNVTSVFQYEGLMRSCVRQSSGF 60
QY 61 TECRPYFTILGLPAMLOAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TECRPYFTILGLPAMLOAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
QY 121 SGIMFIVSGLCIAAGSVFANMLVTNFMWSTANMYTGMGMVQVQVQVQVQVQVQVQV 180
Db 121 SGIMFIVSGLCIAAGSVFANMLVTNFMWSTANMYTGMGMVQVQVQVQVQVQVQVQV 180
QY 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHVSVAIKPGGFKASTGFGSNTKNKKI 240
Db 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHVSVAIKPGGFKASTGFGSNTKNKKI 240
QY 241 YDGGARTEDEVQSPSKHDYV 261
Db 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 8
US-10-063-745-118
; Sequence 118, Application US/10063745
; Publication No. US20040058411A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,745
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 118
; LENGTH: 261
```

```
QY 241 YDGGARTEDEVQSPSKHDYV 261
Db 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 10
US-10-063-512-118
; Sequence 118, Application US/10063512
; Publication No. US20030018183A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,512
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 118
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-512-118

Query Match 100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVAFLLSILGLAGCIAATGMDMWSMTQDLYDNPTSVFQYEGLRWSCVRQSSGF 60
Db 1 MSTTTCQVAFLLSILGLAGCIAATGMDMWSMTQDLYDNPTSVFQYEGLRWSCVRQSSGF 60

QY 61 TCRPYFTILGLPAMLQAVRALMIVGLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TCRPYFTILGLPAMLQAVRALMIVGLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120

QY 121 SGIMFVSGLCALAGSVFANMLVTNFMSTANNYTGMMGVQTVQTRYTFGAALFVGWV 180
Db 121 SGIMFVSGLCALAGSVFANMLVTNFMSTANNYTGMMGVQTVQTRYTFGAALFVGWV 180

QY 181 AGGLTLIGVMMCIACRGLAPEETNYKAVSHASGHSAVKPGGFKASTFGSNTKNKKI 240
Db 181 AGGLTLIGVMMCIACRGLAPEETNYKAVSHASGHSAVKPGGFKASTFGSNTKNKKI 240

QY 241 YDGGARTEDEVQSPSKHDYV 261
Db 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 11
US-10-063-513-118
; Sequence 118, Application US/10063513
; Publication No. US20030018172A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,513
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 118
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-513-118

Query Match 100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVAFLLSILGLAGCIAATGMDMWSMTQDLYDNPTSVFQYEGLRWSCVRQSSGF 60
Db 1 MSTTTCQVAFLLSILGLAGCIAATGMDMWSMTQDLYDNPTSVFQYEGLRWSCVRQSSGF 60

QY 61 TCRPYFTILGLPAMLQAVRALMIVGLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TCRPYFTILGLPAMLQAVRALMIVGLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120

QY 121 SGIMFVSGLCALAGSVFANMLVTNFMSTANNYTGMMGVQTVQTRYTFGAALFVGWV 180
Db 121 SGIMFVSGLCALAGSVFANMLVTNFMSTANNYTGMMGVQTVQTRYTFGAALFVGWV 180

QY 181 AGGLTLIGVMMCIACRGLAPEETNYKAVSHASGHSAVKPGGFKASTFGSNTKNKKI 240
Db 181 AGGLTLIGVMMCIACRGLAPEETNYKAVSHASGHSAVKPGGFKASTFGSNTKNKKI 240

QY 241 YDGGARTEDEVQSPSKHDYV 261
Db 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 12
US-10-063-515-118
; Sequence 118, Application US/10063515
; Publication No. US20030018173A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,515
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 118
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-515-118

Query Match 100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVAFLLSILGLAGCIAATGMDMWSMTQDLYDNPTSVFQYEGLRWSCVRQSSGF 60
Db 1 MSTTTCQVAFLLSILGLAGCIAATGMDMWSMTQDLYDNPTSVFQYEGLRWSCVRQSSGF 60

QY 61 TCRPYFTILGLPAMLQAVRALMIVGLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TCRPYFTILGLPAMLQAVRALMIVGLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120

QY 121 SGIMFVSGLCALAGSVFANMLVTNFMSTANNYTGMMGVQTVQTRYTFGAALFVGWV 180
Db 121 SGIMFVSGLCALAGSVFANMLVTNFMSTANNYTGMMGVQTVQTRYTFGAALFVGWV 180

QY 181 AGGLTLIGVMMCIACRGLAPEETNYKAVSHASGHSAVKPGGFKASTFGSNTKNKKI 240
Db 181 AGGLTLIGVMMCIACRGLAPEETNYKAVSHASGHSAVKPGGFKASTFGSNTKNKKI 240

QY 241 YDGGARTEDEVQSPSKHDYV 261
Db 241 YDGGARTEDEVQSPSKHDYV 261
```

```
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,513
; CURRENT FILING DATE: 2002-05-01
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 118
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-513-118

Query Match 100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVAFLLSILGLAGCIAATGMDMWSMTQDLYDNPTSVFQYEGLRWSCVRQSSGF 60
Db 1 MSTTTCQVAFLLSILGLAGCIAATGMDMWSMTQDLYDNPTSVFQYEGLRWSCVRQSSGF 60

QY 61 TCRPYFTILGLPAMLQAVRALMIVGLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TCRPYFTILGLPAMLQAVRALMIVGLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120

QY 121 SGIMFVSGLCALAGSVFANMLVTNFMSTANNYTGMMGVQTVQTRYTFGAALFVGWV 180
Db 121 SGIMFVSGLCALAGSVFANMLVTNFMSTANNYTGMMGVQTVQTRYTFGAALFVGWV 180

QY 181 AGGLTLIGVMMCIACRGLAPEETNYKAVSHASGHSAVKPGGFKASTFGSNTKNKKI 240
Db 181 AGGLTLIGVMMCIACRGLAPEETNYKAVSHASGHSAVKPGGFKASTFGSNTKNKKI 240

QY 241 YDGGARTEDEVQSPSKHDYV 261
Db 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 12
US-10-063-515-118
; Sequence 118, Application US/10063515
; Publication No. US20030018173A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,515
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 118
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-515-118

Query Match 100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVAFLLSILGLAGCIAATGMDMWSMTQDLYDNPTSVFQYEGLRWSCVRQSSGF 60
Db 1 MSTTTCQVAFLLSILGLAGCIAATGMDMWSMTQDLYDNPTSVFQYEGLRWSCVRQSSGF 60

QY 61 TCRPYFTILGLPAMLQAVRALMIVGLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120
Db 61 TCRPYFTILGLPAMLQAVRALMIVGLGAIGLLVSIFALKCIRIGSMEDSAKANMTLT 120

QY 121 SGIMFVSGLCALAGSVFANMLVTNFMSTANNYTGMMGVQTVQTRYTFGAALFVGWV 180
Db 121 SGIMFVSGLCALAGSVFANMLVTNFMSTANNYTGMMGVQTVQTRYTFGAALFVGWV 180

QY 181 AGGLTLIGVMMCIACRGLAPEETNYKAVSHASGHSAVKPGGFKASTFGSNTKNKKI 240
Db 181 AGGLTLIGVMMCIACRGLAPEETNYKAVSHASGHSAVKPGGFKASTFGSNTKNKKI 240

QY 241 YDGGARTEDEVQSPSKHDYV 261
Db 241 YDGGARTEDEVQSPSKHDYV 261
```

Db 61 TCRPYFTILGLPAMLQAVRALMIVIGVLGAIGLLVSIFALKICIRIGSMEDSAKANMTLT 120
QY 121 SGIMFIVSGLCIAIGSVFANMLVTNFMWSTANMTYTGGMVQVTVQTRYTFGAALFVGWV 180
Db 121 SGIMFIVSGLCIAIGSVFANMLVTNFMWSTANMTYTGGMVQVTVQTRYTFGAALFVGWV 180
QY 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHVSVAIKPGGFKASTGFGSNTKNKKI 240
Db 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHVSVAIKPGGFKASTGFGSNTKNKKI 240
QY 241 YDGGARTEDEVQSPSKHDYV 261
Db 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 13

US-10-063-549-118
; Sequence 118, Application US/10063549
; Publication No. US20030027986A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,549
; CURRENT FILING DATE: 2002-05-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 118
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-549-118

Query Match 100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRSQSGF 60
Db 1 MSTTTCQVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRSQSGF 60
QY 61 TCRPYFTILGLPAMLQAVRALMIVIGVLGAIGLLVSIFALKICIRIGSMEDSAKANMTLT 120
Db 61 TCRPYFTILGLPAMLQAVRALMIVIGVLGAIGLLVSIFALKICIRIGSMEDSAKANMTLT 120
QY 121 SGIMFIVSGLCIAIGSVFANMLVTNFMWSTANMTYTGGMVQVTVQTRYTFGAALFVGWV 180
Db 121 SGIMFIVSGLCIAIGSVFANMLVTNFMWSTANMTYTGGMVQVTVQTRYTFGAALFVGWV 180
QY 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHVSVAIKPGGFKASTGFGSNTKNKKI 240
Db 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHVSVAIKPGGFKASTGFGSNTKNKKI 240
QY 241 YDGGARTEDEVQSPSKHDYV 261
Db 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 14

US-10-063-569-118
; Sequence 118, Application US/10063569
; Publication No. US20030018168A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.

; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,569
; CURRENT FILING DATE: 2002-05-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 118
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-569-118

Query Match 100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTTTCQVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRSQSGF 60
Db 1 MSTTTCQVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRSQSGF 60
QY 61 TCRPYFTILGLPAMLQAVRALMIVIGVLGAIGLLVSIFALKICIRIGSMEDSAKANMTLT 120
Db 61 TCRPYFTILGLPAMLQAVRALMIVIGVLGAIGLLVSIFALKICIRIGSMEDSAKANMTLT 120
QY 121 SGIMFIVSGLCIAIGSVFANMLVTNFMWSTANMTYTGGMVQVTVQTRYTFGAALFVGWV 180
Db 121 SGIMFIVSGLCIAIGSVFANMLVTNFMWSTANMTYTGGMVQVTVQTRYTFGAALFVGWV 180
QY 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHVSVAIKPGGFKASTGFGSNTKNKKI 240
Db 181 AGGLTLIGGVMMCIACRGLAPEETNYKAVSHASGHVSVAIKPGGFKASTGFGSNTKNKKI 240
QY 241 YDGGARTEDEVQSPSKHDYV 261
Db 241 YDGGARTEDEVQSPSKHDYV 261

RESULT 15

US-10-063-551-118
; Sequence 118, Application US/10063551
; Publication No. US20020183494A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,551
; CURRENT FILING DATE: 2002-05-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 118
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-551-118

Query Match 100.0%; Score 1357; DB 12; Length 261;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSTTTCQVAFLLSIILGLAGCIAATGMDMNSTQDLYDNPVTSVFQYEGLRSCVRQSSGF	60
Db	1	MSTTTCQVAFLLSIILGLAGCIAATGMDMNSTQDLYDNPVTSVFQYEGLRSCVRQSSGF	60
QY	61	TECRPYETIILGLPAMLOAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANNTLT	120
Db	61	TECRPYETIILGLPAMLOAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANNTLT	120
QY	121	SGIMFIVSGLCAIAGVSVPANMLVTNFMSTANNYTGMGGMVQTVQTRYTFGAALFVGWV	180
Db	121	SGIMFIVSGLCAIAGVSVPANMLVTNFMSTANNYTGMGGMVQTVQTRYTFGAALFVGWV	180
QY	181	AGGLTLIGGYMNCIACKGLAPEETNYKAVSYHASGHSVAYKPGGFKAFTGFSNTKNKKI	240
Db	181	AGGLTLIGGYMNCIACKGLAPEETNYKAVSYHASGHSVAYKPGGFKAFTGFSNTKNKKI	240
QY	241	YDGGARTEDVQSYPSKHDYV	261
Db	241	YDGGARTEDVQSYPSKHDYV	261

Search completed: September 1, 2004, 16:49:39
Job time : 50 secs

Pred. No. is the number of results predicted by chance to have a

TITLE Secreted and transmembrane polypeptides and nucleic acids encoding

JOURNAL the same

Patent: WO 0116318-A 117 08-MAR-2001;

Genentech, Inc. (US)

FEATURES Location/Qualifiers

1..2121

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 2121; DB 6; Length 2121;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTCCCTCAGAGCGGTTAGCTTCACACCTTCGGCAGCAGAGGGCGGAGCTTCT 60

Db 1 GAGCTCCCTCAGAGCGGTTAGCTTCACACCTTCGGCAGCAGAGGGCGGAGCTTCT 60

QY 61 GCGAGGCGGAGGGCGGCGCCAGGATCATGTCCACCACACATGCCAAGTGTGGCGT 120

Db 61 GCGAGGCGGAGGGCGGCGCCAGGATCATGTCCACCACACATGCCAAGTGTGGCGT 120

QY 121 TCCTCCTGTCATCCTGGGGCTGGCGGCTGCATCGCGGCCACCGGGATGACATGTGGA 180

Db 121 TCCTCCTGTCATCCTGGGGCTGGCGGCTGCATCGCGGCCACCGGGATGACATGTGGA 180

QY 181 GCACCCAGGACCTGTACGACAAACCCCGTCACTCGGTGTTCCAGTACGAGGGCTCTGGA 240

Db 181 GCACCCAGGACCTGTACGACAAACCCCGTCACTCGGTGTTCCAGTACGAGGGCTCTGGA 240

QY 241 GGAGCTGCGTGAGGCGAGGTTCCAGGCTTCACCGAATGAGGCGCTTATTTTCAACCTCTGG 300

Db 241 GGAGCTGCGTGAGGCGAGGTTCCAGGCTTCACCGAATGAGGCGCTTATTTTCAACCTCTGG 300

QY 301 GACTTCAGACCATGCTGAGGCGAGTGGCGGCTGATGATCGTAGGCAATGCGTCTGGGTG 360

Db 301 GACTTCAGACCATGCTGAGGCGAGTGGCGGCTGATGATCGTAGGCAATGCGTCTGGGTG 360

QY 361 CCATTGGGCTCTGGTATCATCTTTGGCCCTGAAATGCAATCGCATTTGGCAGTGGAGG 420

Db 361 CCATTGGGCTCTGGTATCATCTTTGGCCCTGAAATGCAATCGCATTTGGCAGTGGAGG 420

QY 421 ACTCTGCCAAAGCCAAACATGACACTGACCTCGGATCATGTTCTTCTCAGGTCTTT 480

Db 421 ACTCTGCCAAAGCCAAACATGACACTGACCTCGGATCATGTTCTTCTCAGGTCTTT 480

QY 481 GTGCAATTGCTGGAGTGTCTGTTTGGCAACATGCTGGTGACTAACTTCTGGATGTCCA 540

Db 481 GTGCAATTGCTGGAGTGTCTGTTTGGCAACATGCTGGTGACTAACTTCTGGATGTCCA 540

QY 541 CAGCTAACATGTACACCGCATGGTGGGATGGTGAGACTGTTTACAGCCAGGTACACAT 600

Db 541 CAGCTAACATGTACACCGCATGGTGGGATGGTGAGACTGTTTACAGCCAGGTACACAT 600

QY 601 TTGGTGGGCTCTGTTTGGGGCTGGGCTCGCTGGAGGCTTCACTAAATTGGGGGTGGA 660

Db 601 TTGGTGGGCTCTGTTTGGGGCTGGGCTCGCTGGAGGCTTCACTAAATTGGGGGTGGA 660

QY 661 TGAATGTGCATCGCTCGCGGGCTGGCCACAGAGAAACCAACTACAAAGCCGTTTCTT 720

Db 661 TGAATGTGCATCGCTCGCGGGCTGGCCACAGAGAAACCAACTACAAAGCCGTTTCTT 720

QY 721 ATCATGCTCAGGCCACAGTGTTCCTTACAGGCTGGAGCTTCAAGGCCAGCTGGCT 780

Db 721 ATCATGCTCAGGCCACAGTGTTCCTTACAGGCTGGAGCTTCAAGGCCAGCTGGCT 780

QY 781 TTGGGTCCAAACCAAAAACAAGATATACGATGGAGGTGCCCGCACAGAGACGAGG 840

Db 781 TTGGGTCCAAACCAAAAACAAGATATACGATGGAGGTGCCCGCACAGAGACGAGG 840

QY 841 TACAATCTTATCTTCCAAAGCAGTATGTGTAATGCTTAAGACCTCTCAGCAGGGC 900

Db 841 TACAATCTTATCTTCCAAAGCAGTATGTGTAATGCTTAAGACCTCTCAGCAGGGC 900

841 TACAATCTTATCTTCCAAAGCAGTATGTGTAATGCTTAAAGACCTCTCAGCAGGGC 900

QY 901 GGAAGAACTCCCGAGAGCTCACCCAAAACCAAGAGATCCCATCTAGATTTCTTCTT 960

Db 901 GGAAGAACTCCCGAGAGCTCACCCAAAACCAAGAGATCCCATCTAGATTTCTTCTT 960

QY 961 GCTTTTGCATCAGCTGGAGTTAGAAAGCCTCGATTTTCACTTTGGAGAGGCCAAAT 1020

Db 961 GCTTTTGCATCAGCTGGAGTTAGAAAGCCTCGATTTTCACTTTGGAGAGGCCAAAT 1020

QY 1021 GGTCTTAGCCTCAGTCTCTCTTAATATTCACCATAAAAACAGCTGAGTTATTTATG 1080

Db 1021 GGTCTTAGCCTCAGTCTCTCTTAATATTCACCATAAAAACAGCTGAGTTATTTATG 1080

QY 1081 AATTAGAGGCTATAGCTCACATTTTCAATCTCTATTTCTTTTAAATAAATCTTCT 1140

Db 1081 AATTAGAGGCTATAGCTCACATTTTCAATCTCTATTTCTTTTAAATAAATCTTCT 1140

QY 1141 ACTCTGATGAGAGTGTGTTTAACTCTCTCACATTTTGTATGATTTAGACAGACT 1200

Db 1141 ACTCTGATGAGAGTGTGTTTAACTCTCTCACATTTTGTATGATTTAGACAGACT 1200

QY 1201 CCCCTCTTCTCTAGTCAATAAACCCATTTGATGATCTATTTCCAGCTTATCCCAG 1260

Db 1201 CCCCTCTTCTCTAGTCAATAAACCCATTTGATGATCTATTTCCAGCTTATCCCAG 1260

QY 1261 AAAAATTTTGAAGAAAGAGTAGACCCAAAGATGTTATTTTCTGCTGTTGAATTTGT 1320

Db 1261 AAAAATTTTGAAGAAAGAGTAGACCCAAAGATGTTATTTTCTGCTGTTGAATTTGT 1320

QY 1321 CTCCACCCCTTGGCTAGTAAATAACACCTTACTGAAAGAAAGCAATAAGAGAAA 1380

Db 1321 CTCCACCCCTTGGCTAGTAAATAACACCTTACTGAAAGAAAGCAATAAGAGAAA 1380

QY 1381 GATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTGTGATCTTAAAGTT 1440

Db 1381 GATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTGTGATCTTAAAGTT 1440

QY 1441 ACCAAACCAAGTCAATTTTCTAGTTGAGGCAACCAACCTTTCTACTGCTGGACATCT 1500

Db 1441 ACCAAACCAAGTCAATTTTCTAGTTGAGGCAACCAACCTTTCTACTGCTGGACATCT 1500

QY 1501 TCTTATTCAGCAACACCATTTCTAGAGTTTCTCAGCTCTCCTCAGAGTCTCTTCTTCT 1560

Db 1501 TCTTATTCAGCAACACCATTTCTAGAGTTTCTCAGCTCTCCTCAGAGTCTCTTCTTCT 1560

QY 1561 GTGCGGGTCAAGAAATGTCCTTAGATGAATGAGAAATATTTTAAATTTAAGTCTC 1620

Db 1561 GTGCGGGTCAAGAAATGTCCTTAGATGAATGAGAAATATTTTAAATTTAAGTCTC 1620

QY 1621 TAAATATAGTTAAATAAATGTTTAAATGATACTATCTCTGTAATAGC 1680

Db 1621 TAAATATAGTTAAATAAATGTTTAAATGATACTATCTCTGTAATAGC 1680

QY 1681 CTCACCCCTACATGTGATAGAGAAATGAAAAATAATGCTTTGACATTTCTATAT 1740

Db 1681 CTCACCCCTACATGTGATAGAGAAATGAAAAATAATGCTTTGACATTTCTATAT 1740

QY 1741 GGTACTTTGTAAGTCACTTTAAGTACAAATTTCCATGAAAAGCTCACACCTGTATCT 1800

Db 1741 GGTACTTTGTAAGTCACTTTAAGTACAAATTTCCATGAAAAGCTCACACCTGTATCT 1800

QY 1801 AGCATTTGGAGGCTGAGAGAGGATCATCTTGGAGCCAGAGTTTCAGAGCTAGCCTG 1860

Db 1801 AGCATTTGGAGGCTGAGAGAGGATCATCTTGGAGCCAGAGTTTCAGAGCTAGCCTG 1860

QY 1861 GGCACATCGAGAGCCCTGTCTCTACAAAATACAGAGAAATAATCAGCCAGTCATGG 1920

Db 1861 GGCACATCGAGAGCCCTGTCTCTACAAAATACAGAGAAATAATCAGCCAGTCATGG 1920

QY 1921 TGSCATACACTCTGTAGTCCAGCATTTCCGGAGGCTGAGGTGGAGGATCACTTGGAGCC 1980

Db 1921 TGSCATACACTCTGTAGTCCAGCATTTCCGGAGGCTGAGGTGGAGGATCACTTGGAGCC 1980

QY 1981 AGGAGGTTGGGCTGCAGTGAGCCATGATCACACCACTGCACTCCAGCCAGGTGACATA 2040
|||||
Db 1981 AGGAGGTTGGGCTGCAGTGAGCCATGATCACACCACTGCACTCCAGCCAGGTGACATA 2040
QY 2041 GCGAGATCCTGTCTAAAAAATAAATAAATAAATGGAACACAGCAAGTCTTAGGAAGT 2100
|||||
Db 2041 GCGAGATCCTGTCTAAAAAATAAATAAATAAATGGAACACAGCAAGTCTTAGGAAGT 2100
QY 2101 AGGTTAAACCTAAATCTTTAA 2121
|||||
Db 2101 AGGTTAAACCTAAATCTTTAA 2121

RESULT 2
AX376288
LOCUS 2121 bp DNA linear PAT 01-MAR-2002
DEFINITION Sequence 355 from Patent WO0168848.
ACCESSION AX376288
VERSION AX376288.1 GI:19170535
KEYWORDS Homo sapiens (human)
SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ORGANISM Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS 1 Baker, K.P., Chen, J., Desnoyers, L., Goddard, A., Godowski, P.J.,
Gurney, A.L., Pan, J., Smith, V., Watanabe, C.K., Wood, W.I. and
Zhang, Z.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0168848-A 355 20-SEP-2001;
Genentech, Inc. (US)
FEATURES
source Location/Qualifiers
1. . 2121
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 2121; DB 6; Length 2121;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGTCCCTCAGGAGCGGTTAGCTTACACCTTCGGCAGCAGGAGGCGGCGAGCTTCT 60
|||||
Db 1 GAGTCCCTCAGGAGCGGTTAGCTTACACCTTCGGCAGCAGGAGGCGGCGAGCTTCT 60
QY 61 CGCAGCGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
|||||
Db 61 CGCAGCGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
QY 121 TCCTCTGTCATCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
|||||
Db 121 TCCTCTGTCATCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
QY 181 GCACCCAGGACCTGTACGCAACACCCCGTCACCTCCGTTTCCAGTACGAGGCGCTCGA 240
|||||
Db 181 GCACCCAGGACCTGTACGCAACACCCCGTCACCTCCGTTTCCAGTACGAGGCGCTCGA 240
QY 241 GGAGCTGGGTGAGGCGAGTTGAGGCTTCCGAAATGCAATGCAATGCAATGCAATGCA 300
|||||
Db 241 GGAGCTGGGTGAGGCGAGTTGAGGCTTCCGAAATGCAATGCAATGCAATGCAATGCA 300
QY 301 GACTTCCAGCCATGTGAGGCGAGTTCAGGCTTCCGAAATGCAATGCAATGCAATGCA 360
|||||
Db 301 GACTTCCAGCCATGTGAGGCGAGTTCAGGCTTCCGAAATGCAATGCAATGCAATGCA 360
QY 361 CCATTGGGCTCTGGTATCCATCTTTGGCCCTGAAATGCAATGCAATGCAATGCAATG 420
|||||
Db 361 CCATTGGGCTCTGGTATCCATCTTTGGCCCTGAAATGCAATGCAATGCAATGCAATG 420
QY 421 ACTCTGCCAAGCCAAACATGACACTGACCTCCGGGATCATGTTCTCAGGTCTTT 480
|||||

Db 421 ACTCTGCCAAGCCAAACATGACACTGACCTCCGGATCATGTTCAATGCTCAGGTCTTT 480
QY 481 GTGCAATGTCTGGAGTGTCTGTGTTGCAACACTGCTGGTGACTAACTTCTGGATGTCCA 540
|||||
Db 481 GTGCAATGTCTGGAGTGTCTGTGTTGCAACACTGCTGGTGACTAACTTCTGGATGTCCA 540
QY 541 CAGCTAACATGTACACCGGATGCGGTGGGATGCTGGTGGAGCTGTTTCAAGCCAGGTACAT 600
|||||
Db 541 CAGCTAACATGTACACCGGATGCGGTGGGATGCTGGTGGAGCTGTTTCAAGCCAGGTACAT 600
QY 601 TTGGTGGGCTCTGTTCTGGTGGGCTGGTGGGCTGGTGGGCTGGTGGGCTGGTGGG 660
|||||
Db 601 TTGGTGGGCTCTGTTCTGGTGGGCTGGTGGGCTGGTGGGCTGGTGGGCTGGTGGG 660
QY 661 TGATGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
|||||
Db 661 TGATGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 ATCATGCTCAGGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
|||||
Db 721 ATCATGCTCAGGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 781 TTGGGTCCAAACACCAAAACCAAGAGATATACGATGGAGGTGCCCGCACAGAGACAGG 840
|||||
Db 781 TTGGGTCCAAACACCAAAACCAAGAGATATACGATGGAGGTGCCCGCACAGAGACAGG 840
QY 841 TACAATCTTATCTTCCAGCAGCAGTATGCTGTAATGCTCTAAGACCTCTCAGCAGCGGC 900
|||||
Db 841 TACAATCTTATCTTCCAGCAGCAGTATGCTGTAATGCTCTAAGACCTCTCAGCAGCGGC 900
QY 901 GGAAGAACTCCCGGAGAGCTCACCACCAAAACCAAGAGATCCCATCTAGATTTCTTT 960
|||||
Db 901 GGAAGAACTCCCGGAGAGCTCACCACCAAAACCAAGAGATCCCATCTAGATTTCTTT 960
QY 961 GCTTTTGACTCAGCTGGAAGTTAGAAAGCCTCGATTTTCACTTTTGGAGAGCCAAAT 1020
|||||
Db 961 GCTTTTGACTCAGCTGGAAGTTAGAAAGCCTCGATTTTCACTTTTGGAGAGCCAAAT 1020
QY 1021 GGTCTTAGCTCAGTCTCTGCTCTAAATATTCACATAAAACAGCTGAGTTATTTATG 1080
|||||
Db 1021 GGTCTTAGCTCAGTCTCTGCTCTAAATATTCACATAAAACAGCTGAGTTATTTATG 1080
QY 1081 AATTAGAGCTATAGCTCAGTTTCAATTCCTCTATTTCTTTTAAATATACTTTCT 1140
|||||
Db 1081 AATTAGAGCTATAGCTCAGTTTCAATTCCTCTATTTCTTTTAAATATACTTTCT 1140
QY 1141 ACTCTGATGAGAGAAATGTTGTTTAAATCTCTCTCACAATTTTGTGATTTAGACAGCT 1200
|||||
Db 1141 ACTCTGATGAGAGAAATGTTGTTTAAATCTCTCTCACAATTTTGTGATTTAGACAGCT 1200
QY 1201 CCCCCTCTTCTCTCTAGTCAATTAACCCATTTGATGATCTATTTCCAGCTTATCCCCAAG 1260
|||||
Db 1201 CCCCCTCTTCTCTCTAGTCAATTAACCCATTTGATGATCTATTTCCAGCTTATCCCCAAG 1260
QY 1261 AAAAATTTTGAAGAAAGTAGAGCCCAAGATGTTATTTCTGCTGTTTGAATTTTCT 1320
|||||
Db 1261 AAAAATTTTGAAGAAAGTAGAGCCCAAGATGTTATTTCTGCTGTTTGAATTTTCT 1320
QY 1321 CTCCTCCACCCCACTTGGCTAGTAAATAACACTTACTGAAAGAAAGCAATAAGAGAA 1380
|||||
Db 1321 CTCCTCCACCCCACTTGGCTAGTAAATAACACTTACTGAAAGAAAGCAATAAGAGAA 1380
QY 1381 GATATTTGATATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTGATCTTAAAGTT 1440
|||||
Db 1381 GATATTTGATATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTGATCTTAAAGTT 1440
QY 1441 ACCAAACCAAGTCAATTTTCTAGTTTGGAGCAACCAACTTTCTACTGCTGCTGACATCT 1500
|||||
Db 1441 ACCAAACCAAGTCAATTTTCTAGTTTGGAGCAACCAACTTTCTACTGCTGCTGACATCT 1500
QY 1501 TCTTATTACAGCAACCACTTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCTCTTTCT 1560
|||||

Db 1501 TCATTATACAGCAACACCATTTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCTCTTTCT 1560
Qy 1561 GTCGGGCTCAGAAATTTGTCCTAGATGAATGAGAAATTTATTTTAAATTTAAAGTCC 1620
Db 1561 GTCGGGCTCAGAAATTTGTCCTAGATGAATGAGAAATTTATTTTAAATTTAAAGTCC 1620
Qy 1621 TAAATATAGTTTAAATTAATTAATTTTAAATTAATTTTAAATTTAAAGTCC 1680
Db 1621 TAAATATAGTTTAAATTAATTAATTTTAAATTAATTTTAAATTTAAAGTCC 1680
Qy 1681 CTCACCCCTCATGATGGATAGAGAAATGAGAAATTAATTTGCTTGAATGAGTCTATAT 1740
Db 1681 CTCACCCCTCATGATGGATAGAGAAATGAGAAATTAATTTGCTTGAATGAGTCTATAT 1740
Qy 1741 GGTACTTTGTAAGTCTAGTCTTAAATTTTAAATTTTAAATTTTAAATTTTAAAGTCC 1800
Db 1741 GGTACTTTGTAAGTCTAGTCTTAAATTTTAAATTTTAAATTTTAAATTTTAAAGTCC 1800
Qy 1801 AGCACTTTGGAGGCTGAGGAGGAGGATCACTTTGAGCCCAAGAGTTTCGAGACTAGCCTG 1860
Db 1801 AGCACTTTGGAGGCTGAGGAGGAGGATCACTTTGAGCCCAAGAGTTTCGAGACTAGCCTG 1860
Qy 1861 GGCACATGAGAGAGCCCTGCTCTCAAAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
Db 1861 GGCACATGAGAGAGCCCTGCTCTCAAAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
Qy 1921 TGGCATACACCTGTAGTCCAGCATTCCGGAGGCTGAGGTGGAGGAGTCACTTTGAGCCC 1980
Db 1921 TGGCATACACCTGTAGTCCAGCATTCCGGAGGCTGAGGTGGAGGAGTCACTTTGAGCCC 1980
Qy 1981 AGGAGGTTGGGCTGAGTGGAGGATGATCACCACTGCACTCCAGCCAGGTGACATA 2040
Db 1981 AGGAGGTTGGGCTGAGTGGAGGATGATCACCACTGCACTCCAGCCAGGTGACATA 2040
Qy 2041 GCGAGATCTGTCTAAAAAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 2100
Db 2041 GCGAGATCTGTCTAAAAAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 2100
Qy 2101 AGGTTAAACTAAATTTCTTAA 2121
Db 2101 AGGTTAAACTAAATTTCTTAA 2121

RESULT 3
AX697257 2121 bp DNA linear PAT 02-APR-2003
LOCUS Sequence 325 from Patent WO0078961.
DEFINITION AX697257
ACCESSION AX697257.1 GI:29498417
VERSION
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Ferrara, N., Stewart, T.A., Williams, P.M., Baker, K.P., Desnoyers, L., Eaton, D.L., Gao, W.Q., Pan, J., Botstein, D., Fong, S., Goddard, A., Godowski, P.J., Gurney, A.L., Smith, V., Tamas, D., Wood, W.I., Grimaldi, C.J., Hillan, K.J., Paoni, N.F., Roy, M.A., and Watanabe, C.K.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding the same
JOURNAL Patent: WO 0078961-A 325 28-DEC-2000;
Genentech Inc. (US)
FEATURES
source Location/Qualifiers
1..2121
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 2121; DB 6; Length 2121;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTCTCCCTCAGAGGGCGGTAGCTTTCACCTTCCGAGCAGAGGGCGGAGCTTCT 60
Db 1 GAGTCTCCCTCAGAGGGCGGTAGCTTTCACCTTCCGAGCAGAGGGCGGAGCTTCT 60
Qy 61 CGCAGGGCGGAGGGCGGCGGAGGATCATGTCCACACACATGTCACCAAGTGGTGGCGT 120
Db 61 CGCAGGGCGGAGGGCGGCGGAGGATCATGTCCACACACATGTCACCAAGTGGTGGCGT 120
Qy 121 TCCTCTCTGTCATCTCGGGGCTGCGCGGCTGCAATCGCGGCCACCGGATGACATGTGGA 180
Db 121 TCCTCTCTGTCATCTCGGGGCTGCGCGGCTGCAATCGCGGCCACCGGATGACATGTGGA 180
Qy 181 GCACCCAGGACCTGTAGCAAAACCCCTGACCTCCGTTCCAGTACGAGGGCTCTGGA 240
Db 181 GCACCCAGGACCTGTAGCAAAACCCCTGACCTCCGTTCCAGTACGAGGGCTCTGGA 240
Qy 241 GGAGCTCGGTGAGGACAGTTCAGGCTTCCAGGATGAGGCGCTATTTCCATCCCTCTG 300
Db 241 GGAGCTCGGTGAGGACAGTTCAGGCTTCCAGGATGAGGCGCTATTTCCATCCCTCTG 300
Qy 301 GACTTCCAGCCATCTGAGGACAGTGGAGCCCTGATGATGATGATGATGATGATGATGAT 360
Db 301 GACTTCCAGCCATCTGAGGACAGTGGAGCCCTGATGATGATGATGATGATGATGATGAT 360
Qy 361 CCATTGGCTCTCTGATTCATCTTCCCTGAAATGATGATGATGATGATGATGATGATGAT 420
Db 361 CCATTGGCTCTCTGATTCATCTTCCCTGAAATGATGATGATGATGATGATGATGATGAT 420
Qy 421 ACTCTGCCAAAGCCAAACATGACACTGCTCCGGGATCATGTTCTCTCAGTCTTTT 480
Db 421 ACTCTGCCAAAGCCAAACATGACACTGCTCCGGGATCATGTTCTCTCAGTCTTTT 480
Qy 481 GTGCAATTGCTGGAGTGTCTGTTTGGCAATGCTGGTGAATGCTGATGATGATGATGAT 540
Db 481 GTGCAATTGCTGGAGTGTCTGTTTGGCAATGCTGGTGAATGCTGATGATGATGATGAT 540
Qy 541 CAGCTAAATGCTACACCCGATGCTGGTGGATGCTGAGTGGTGGTGGTGGTGGTGGTGG 600
Db 541 CAGCTAAATGCTACACCCGATGCTGGTGGATGCTGAGTGGTGGTGGTGGTGGTGGTGG 600
Qy 601 TTGGTGGGCTCTGTTTGGGCTGGTGGGCTGGTGGGCTGGTGGGCTGGTGGGCTGGTGG 660
Db 601 TTGGTGGGCTCTGTTTGGGCTGGTGGGCTGGTGGGCTGGTGGGCTGGTGGGCTGGTGG 660
Qy 661 TGATGTGATCGCTTCCCGGGGCTGGCAACAGAGAGATATAGTGGAGGCTCCCGCAGAGG 720
Db 661 TGATGTGATCGCTTCCCGGGGCTGGCAACAGAGAGATATAGTGGAGGCTCCCGCAGAGG 720
Qy 721 ATCATGCTCAGGCGCACAGTGTGGCTCAAGCTGGAGGCTTCAAGGCGAGCAGCTGGCT 780
Db 721 ATCATGCTCAGGCGCACAGTGTGGCTCAAGCTGGAGGCTTCAAGGCGAGCAGCTGGCT 780
Qy 781 TTGGGTCCAAACACCAAGAGATATAGTGGAGGCTCCCGCAGAGGAGGAGGAGGAGG 840
Db 781 TTGGGTCCAAACACCAAGAGATATAGTGGAGGCTCCCGCAGAGGAGGAGGAGGAGG 840
Qy 841 TACAATCTTTATCTTCCAAAGCAGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 900
Db 841 TACAATCTTTATCTTCCAAAGCAGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 900
Qy 901 GGAAGAACTCCCGGAGGCTCACCCAAAGAGATCCCAAGAGATCCCAAGAGATTTCTTCT 960
Db 901 GGAAGAACTCCCGGAGGCTCACCCAAAGAGATCCCAAGAGATCCCAAGAGATTTCTTCT 960
Qy 961 GCTTTTGAATCTCAGCTGGAAGTTAGAAAAGCTTCGATTTCTATTTCTTTGAGAGGCCA 1020
Db 961 GCTTTTGAATCTCAGCTGGAAGTTAGAAAAGCTTCGATTTCTATTTCTTTGAGAGGCCA 1020
Qy 1021 GGTCTTAGGCTCAGTCTCTGCTCTAAATTTCCCAATTAATTTCCCAATTAATTTATG 1080
Db 1021 GGTCTTAGGCTCAGTCTCTGCTCTAAATTTCCCAATTAATTTCCCAATTAATTTATG 1080

```
QY 1081 AATTAGAGGCTATAGCTCACATTTTCAATCTCTATTTCTTTTAAATAAATCTTCT 1140
Db 1081 AATTAGAGGCTATAGCTCACATTTTCAATCTCTATTTCTTTTAAATAAATCTTCT 1140
QY 1141 ACTCTGATGAGAGATGGTGTAAATCTCTCTCTACATTTTGATGATTAGACAGCT 1200
Db 1141 ACTCTGATGAGAGATGGTGTAAATCTCTCTCTACATTTTGATGATTAGACAGCT 1200
QY 1201 CCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
Db 1201 CCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
QY 1261 AAAATCTTTGAAGGAAGAGTAGACCCAAAGATGTTTATTTTCGTCTTTGAATTTGT 1320
Db 1261 AAAATCTTTGAAGGAAGAGTAGACCCAAAGATGTTTATTTTCGTCTTTGAATTTGT 1320
QY 1321 CTCCCAACCCCACTCTGGCTAGTAAATAACACTTACTGAAGAGAGCAATAGAGAAA 1380
Db 1321 CTCCCAACCCCACTCTGGCTAGTAAATAACACTTACTGAAGAGAGCAATAGAGAAA 1380
QY 1381 GATATTTGTAATCTCTCCAGCCATGATCTCGGTTTTTCTTACACTGTGATCTTAAAGTT 1440
Db 1381 GATATTTGTAATCTCTCCAGCCATGATCTCGGTTTTTCTTACACTGTGATCTTAAAGTT 1440
QY 1441 ACCAAACCAAGTCAATTTTCTGAGGCAACCAACCTTTCTACTGCTGTTGACATCT 1500
Db 1441 ACCAAACCAAGTCAATTTTCTGAGGCAACCAACCTTTCTACTGCTGTTGACATCT 1500
QY 1501 TCTTATTACAGCAACCACTCTAGGATTTCTGAGCTCTCTGAGTCTCTCTTTCT 1560
Db 1501 TCTTATTACAGCAACCACTCTAGGATTTCTGAGCTCTCTGAGTCTCTCTTTCT 1560
QY 1561 GTCGCGGCTCAGAAATGTCCTTAGATGAATGAGAAATATTTTAAATTTAGTCC 1620
Db 1561 GTCGCGGCTCAGAAATGTCCTTAGATGAATGAGAAATATTTTAAATTTAGTCC 1620
QY 1621 TAAATATAGTTAAATTAATATGTTTTAGTAAATGATACATCTCTGTGAATAGC 1680
Db 1621 TAAATATAGTTAAATTAATATGTTTTAGTAAATGATACATCTCTGTGAATAGC 1680
QY 1681 CTCACCCCTACATGCGATAGAGGAAATGAAAAATATGTTTGAATGCTCTATAT 1740
Db 1681 CTCACCCCTACATGCGATAGAGGAAATGAAAAATATGTTTGAATGCTCTATAT 1740
QY 1741 GGTACTTTGTAAGTCAATGCTTAAGTACAAATTCATGAAAGCTCACACCTGTAATCT 1800
Db 1741 GGTACTTTGTAAGTCAATGCTTAAGTACAAATTCATGAAAGCTCACACCTGTAATCT 1800
QY 1801 AGCATTCTGGAGGCTGAGAGAGGATCACTTGAGCCAGAGTTGAGAGTACCTGCTG 1860
Db 1801 AGCATTCTGGAGGCTGAGAGAGGATCACTTGAGCCAGAGTTGAGAGTACCTGCTG 1860
QY 1861 GGCACATGAGAGCCCTGCTCTACAAATACAGAGAGAAATACAGCAGTCAATGG 1920
Db 1861 GGCACATGAGAGCCCTGCTCTACAAATACAGAGAGAAATACAGCAGTCAATGG 1920
QY 1921 TGGCATACACCTGTAGTCCAGCATTTCCGGAGGCTGAGGTGCGAGATCACTTGAGCCC 1980
Db 1921 TGGCATACACCTGTAGTCCAGCATTTCCGGAGGCTGAGGTGCGAGATCACTTGAGCCC 1980
QY 1981 AGGAGGTTGGGGCTGAGTGAAGCATGATCACACCTGCACTCCAGCCAGGTGACATA 2040
Db 1981 AGGAGGTTGGGGCTGAGTGAAGCATGATCACACCTGCACTCCAGCCAGGTGACATA 2040
QY 2041 GGGAGATCTGCTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2100
Db 2041 GGGAGATCTGCTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2100
QY 2101 AGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2121
Db 2101 AGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2121
```

RESULT 4

AY358479

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

ORIGIN

Query Match

Best Local Similarity

Matches 2121;

Conservative

0;

Mismatches

0;

Indels

0;

Gaps

0;

Length

2121;

DB

9;

Score

100.0%;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

2121;

AY358479 2121 bp mRNA linear PRI 03-OCT-2003

Homo sapiens clone DNA73734 CLDN18 (UNQ778) mRNA, complete cds.

AY358479

AY358479.1 GI:37182080

FLI CDNA.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 2121)

Clark, H.F., Gurney, A.L., Abaya, E., Baker, K., Baldwin, D., Brush, J.,

Chen, J., Chow, B., Chui, C., Crowley, C., Currell, B., Deuel, B.,

Dowd, P., Eaton, D., Foster, J., Grimaldi, C., Gu, Q., Hass, P.E.,

Heldens, S., Huang, A., Kim, H.S., Klimowski, L., Jin, Y., Johnson, S.,

Lee, J., Lewis, L., Liao, D., Mark, M., Robbie, E., Sanchez, C.,

Schoenfeld, J., Seehagiri, S., Simmons, L., Singh, J., Smith, V.,

Stinson, J., Vagts, A., Vandlen, R., Watanabe, C., Wleand, D., Woods, K.,

Xie, M.H., Yansura, D., Yi, S., Yu, G., Yuan, J., Zhang, M., Zhang, Z.,

Goddard, A., Wood, W.I. and Godowski, P.

The Secreted Protein Discovery Initiative (SPDI), a Large-Scale

Effort to Identify Novel Human Secreted and Transmembrane Proteins:

A Bioinformatics Assessment

Genome Res. 13 (10), 2265-2270 (2003)

12975309

2 (bases 1 to 2121)

Clark, H.F.

Direct Submission

Submitted (01-AUG-2003) Department of Bioinformatics, Genentech,

Inc., 1 DNA Way, South San Francisco, CA 94080, USA

Location/Qualifiers

1. .2121

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DNA73734"

1. .2121

/locus_tag="UNQ778"

90. .875

/locus_tag="UNQ778"

/note="PRO1572"

/codon_start=1

/product="CLDN18"

/protein_id="AAQ88843.1"

/db_xref="GI:37182081"

/translation="MSFTTCQVAFLLSLGLAGCAATGMDMSTQDLYDNPVTSVF

QVEGLWRSVCVROSQGTFCRPFYTLGLPAMLOAVRALMIVGLGAILLYSIFALK

CIRGSMEDSAKANNTLPSGIMFIVSGLCALAGVSFVFNMLVTFNFMSTANMTYMG

MYQTQRYTFGAALFVGVAGGLTLIGVMMCIACRGLAPEETNYKAVSYHASGHSV

ATKPGFKASTGFGSNTKNTKKIYDGGARTEDDEVQSPKHDYV"

QY 1 GAGCTCCCTCAGGAGCGGTTAGCTTACACCTTCCGAGCAGGAGCGGCGGCTCT 60

Db 1 GAGCTCCCTCAGGAGCGGTTAGCTTACACCTTCCGAGCAGGAGCGGCGGCTCT 60

QY 61 CGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 120

Db 61 CGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 120

QY 121 TCCTCTCTGTCATCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 180

Db 121 TCCTCTCTGTCATCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 180

QY 181 GCACCCAGGACCTGTACGACACACCCCGCTCAGCTCCGCTTCCAGTACCAAGGCTCTGGA 240

Db 181 GCACCCAGGACCTGTACGACACACCCCGCTCAGCTCCGCTTCCAGTACCAAGGCTCTGGA 240

241 GAGCTCGCTGAGGAGGTTCCAGGCTTCAACGAATGAGGCGCCCTATTTCCACATCTCG 300
 241 GAGCTCGCTGAGGAGGTTCCAGGCTTCAACGAATGAGGCGCCCTATTTCCACATCTCG 300
 301 GACTTCCAGCCATGCTGCTGAGGAGTGGAGCCCTGATGATGCTGATGCTGCTGCTGCTG 360
 301 GACTTCCAGCCATGCTGCTGAGGAGTGGAGCCCTGATGATGCTGATGCTGCTGCTGCTG 360
 361 CCATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 361 CCATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 421 ACTCTGCAAGCCAAACATGACATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 421 ACTCTGCAAGCCAAACATGACATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 481 GTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 481 GTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 541 CAGCTAACATGATACACCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 541 CAGCTAACATGATACACCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 601 TTGGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 601 TTGGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 661 TGATGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 661 TGATGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 721 ATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 721 ATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 781 TTGGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 781 TTGGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 841 TACAATCTTATCTTCCAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 841 TACAATCTTATCTTCCAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 901 GGAAGAACTCCCGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 901 GGAAGAACTCCCGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 961 GCTTTTGACTCAGCTGGAAGTTAGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 961 GCTTTTGACTCAGCTGGAAGTTAGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 1021 GGTCTTAGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 1021 GGTCTTAGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 1081 RAATAGGCTATAGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 1081 RAATAGGCTATAGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 1141 ACTCTGATGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 1141 ACTCTGATGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 1201 CCCCCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 1201 CCCCCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 1261 AAAAATCTTTGAAAGGAGAGTAGAGCCAAAGATGTTATTTCTGCTGCTGCTGCTGCTGCT 1320
 1261 AAAAATCTTTGAAAGGAGAGTAGAGCCAAAGATGTTATTTCTGCTGCTGCTGCTGCTGCT 1320
 1321 CTCCCCACCCCAACTTTGGCTAGTAAATAACATCTTACTGAGGAGAGCAATAAGAGAA 1380

1321 CTCCCCACCCCAACTTTGGCTAGTAAATAACATCTTACTGAGGAGAGCAATAAGAGAA 1380
 1381 GATATTTGTAATCTCTCCAGCCCATGATTCGGTTTCTTACACTGTGATCTTAAAGATT 1440
 1381 GATATTTGTAATCTCTCCAGCCCATGATTCGGTTTCTTACACTGTGATCTTAAAGATT 1440
 1441 ACCAAACCAAGTCATTTTCAGTTTGGAGCAACCAACCTTTCTACTGCTGCTGCTGCTGCT 1500
 1441 ACCAAACCAAGTCATTTTCAGTTTGGAGCAACCAACCTTTCTACTGCTGCTGCTGCTGCT 1500
 1501 TCTTATTACAGCAACACATTTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCTCTTTCT 1560
 1501 TCTTATTACAGCAACACATTTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCTCTTTCT 1560
 1561 GTCCGGGTGAGAAATTTGCTTCTAGATGAATGAGAAATTAATTTTAAATTTAAAGTCC 1620
 1561 GTCCGGGTGAGAAATTTGCTTCTAGATGAATGAGAAATTAATTTTAAATTTAAAGTCC 1620
 1621 TAAATATATGTTAAATTAATTAATTTTCTGAGTAAATGATACACTATCTCTGTAATAGC 1680
 1621 TAAATATATGTTAAATTAATTAATTTTCTGAGTAAATGATACACTATCTCTGTAATAGC 1680
 1681 CTCACCTCTACATCTGATGAGAAATGAGAAATTAATTTGCTTGTGATTTGCTATAT 1740
 1681 CTCACCTCTACATCTGATGAGAAATGAGAAATTAATTTGCTTGTGATTTGCTATAT 1740
 1741 GGTACTTTGTAAGTCATGCTTAAAGTACAAATTCATGAAAGCTCACACCTGTAATCT 1800
 1741 GGTACTTTGTAAGTCATGCTTAAAGTACAAATTCATGAAAGCTCACACCTGTAATCT 1800
 1801 AGCACTTTGGAGGCTGAGGAGGAGGATCACTTGTAGCCAGAGAGTTTCCAGACTAGCCTG 1860
 1801 AGCACTTTGGAGGCTGAGGAGGAGGATCACTTGTAGCCAGAGAGTTTCCAGACTAGCCTG 1860
 1861 G3CAACATG3AGAGAGCCCTGCTCTACAAATACAGAGAGAGAAATCAGCCAGTCATGG 1920
 1861 G3CAACATG3AGAGAGCCCTGCTCTACAAATACAGAGAGAGAAATCAGCCAGTCATGG 1920
 1921 TGGCATACACCTGTAGTCCAGCATTCGGAGGCTGAGTGGGAGGATCACTTGGAGCC 1980
 1921 TGGCATACACCTGTAGTCCAGCATTCGGAGGCTGAGTGGGAGGATCACTTGGAGCC 1980
 1981 AGGAGGTTGGGCTGCTGAGTGGAGGATCACTGATCAGCACTGCTGCTGCTGCTGCTGCT 2040
 1981 AGGAGGTTGGGCTGCTGAGTGGAGGATCACTGATCAGCACTGCTGCTGCTGCTGCTGCT 2040
 2041 GCGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
 2041 GCGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
 2101 AGGTTAAACCTAATTTCTTTAA 2121
 2101 AGGTTAAACCTAATTTCTTTAA 2121

RESULT 5

AK098474

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AK098474 3344 bp mRNA linear PRI 12-SEP-2003
 Homo sapiens cDNA FLJ25608 fis, clone STW00359, highly similar to
 CLAUDIN-18.
 AK098474
 AK098474.1 GI:21758493
 oligo capping; fis (full insert sequence).
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1
 Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
 Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,
 Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,
 Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S., Morinaga, M.,

Db 1259 TTTTGAAGGAAAGAGTAGACCCAAAGATGTTATTTCTGCTGTTGAATTTCTCTCCC 1318
 Qy 1326 CACCCCAACTTGCTAGTAACTAATAACCTTACTGAAGAAGCAATAAGAGAAGATAT 1385
 Db 1319 CACCCCAACTTGCTAGTAACTAATAACCTTACTGAAGAAGCAATAAGAGAAGATAT 1378
 Qy 1386 TTGTAATCTCTCCAGCCCATGATCTCGGTTTCTTACACCTGTGATCTTAAAGTTACCA 1445
 Db 1379 TTGTAATCTCTCCAGCCCATGATCTCGGTTTCTTACACCTGTGATCTTAAAGTTACCA 1438
 Qy 1446 ACCAAAGTCATTTTCAGTTTGAGGCAACCAACCTTCTACTGCTGTGTGACATCTCTTA 1505
 Db 1439 ACCAAAGTCATTTTCAGTTTGAGGCAACCAACCTTCTACTGCTGTGTGACATCTCTTA 1498
 Qy 1506 TTACAGCAACACCACTTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCTCTTCTGTCGC 1565
 Db 1499 TTACAGCAACACCACTTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCTCTTCTGTCGC 1558
 Qy 1566 GGGTCGAAATGTCCTAGATGAATGAGAAATTAATTTTTTAAATTAAGTCCTAAAT 1625
 Db 1559 GGGTCGAAATGTCCTAGATGAATGAGAAATTAATTTTTTAAATTAAGTCCTAAAT 1618
 Qy 1626 ATAGTTAAATTAATATGTTTGTAGTAAATGATACACTATCTCTGGAATAGCCTCAC 1685
 Db 1619 ATAGTTAAATTAATATGTTTGTAGTAAATGATACACTATCTCTGGAATAGCCTCAC 1678
 Qy 1686 CCCTACATGTCGATAGAGAAATGAAATTAATTTCTTTGACATGCTATATGTCAC 1745
 Db 1679 CCCTACATGTCGATAGAGAAATGAAATTAATTTCTTTGACATGCTATATGTCAC 1738
 Qy 1746 TTTGTAAGTCATGCTTAAAGTACAAATTCATGAAAGCTCAC 1788
 Db 1739 TTTGTAAGTCATGCTTAAAGTACAAATTCATGAAAGCTCAC 1781

RESULT 7
 LOCUS BC063845 2040 bp mRNA linear PRI 12-DEC-2003
 DEFINITION Homo sapiens cDNA clone IMAGE:6177081, partial cds.
 ACCESSION BC063845
 VERSION BC063845.1 GI:39795429
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2040)
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
 Carninci,P., Prange,C., Raja,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
 Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalusz,D.E.,
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 2040)
 Strausberg,R.
 Direct Submission

ORIGIN
 Query Match 78.6%; Score 1667; DB 9; Length 2040;
 Best Local Similarity 89.2%; Pred. No. 0;
 Matches 1914; Conservative 0; Mismatches 10; Indels 221; Gaps 2;
 QY 80 GGCCAGGATCATGTCCACCACACATGCCAAGTGTGGGCTTCTCTGTCCATCTCTGG 139
 Db 1 GGCCAGGATCATGTCCACCACACATGCCAAGTGTGGGCTTCTCTGTCCATCTCTGG 60
 QY 140 GCTGGCCGGCTGCATCGCGGCCACCGGGATGACATGTGGAGCAGCCAGGACCTGTACGA 199
 Db 61 GCTGGCCGGCTGCATCGCGGCCACCGGGATGACATGTGGAGCAGCCAGGACCTGTACGA 120
 QY 200 CAAACCCGTCACCTCGGTGTTCCAGTACGAAGGGCTCTGGAGAGCTGGTGGAGCAGAG 259
 Db 121 CAAACCCGTCACCTCGGTGTTCCAGTACGAAGGGCTCTGGAGAGCTGGTGGAGCAGAG 180
 QY 260 TTCAGGCTTCACGATGAGCCCTTATTTACCATCTCTGGAGCTTCAGCCATGCTGCA 319
 Db 181 TTCAGGCTTCACGATGAGCCCTTATTTACCATCTCTGGAGCTTCAGCCATGCTGCA 240
 QY 320 GGCAGTGGAGCCCTGATGATCGTAGGATCGTCTGGTGGTCCATGGCCTCTGTATC 379
 Db 241 GGCAGTGGAGCCCTGATGATCGTAGGATCGTCTGGTGGTCCATGGCCTCTGTATC 300
 QY 380 CATCTTTCCCTGAAATGCATCCGATTCGGAGCATGGAGGACTCTGCCAAAGCCAAAT 439
 Db 301 CATCTTTCCCTGAAATGCATCCGATTCGGAGCATGGAGGACTCTGCCAAAGCCAAAT 360
 QY 440 GACACTGACCTCCGGATCATGTTCAATGTCAGGTCTTTGTGCAATTTGTGGAGTCTC 499
 Db 361 GACACTGACCTCCGGATCATGTTCAATGTCAGGTCTTTGTGCAATTTGTGGAGTCTC 392
 QY 500 TGTGTTTCCCAACATGCTGGTGAACCTTCTGGATGTCACAGCTAACATGTACACCGG 559

JOURNAL

REMARK
COMMENT

Submitted (08-DEC-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Liisa Prabhun, Parvaneh Saedi, Jacqueline
 Schein, Duane Smalusz, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 141 Row: a Column: 13
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 7705960.

FEATURES
source

Location/Qualifiers
 1..2040
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6177081"
 /tissue_type="Peripheral Nervous System, dorsal root
 ganglion"
 /clone_lib="Lupski_dorsal_root_ganglion"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 78.6%; Score 1667; DB 9; Length 2040;
 Best Local Similarity 89.2%; Pred. No. 0;
 Matches 1914; Conservative 0; Mismatches 10; Indels 221; Gaps 2;
 QY 80 GGCCAGGATCATGTCCACCACACATGCCAAGTGTGGGCTTCTCTGTCCATCTCTGG 139
 Db 1 GGCCAGGATCATGTCCACCACACATGCCAAGTGTGGGCTTCTCTGTCCATCTCTGG 60
 QY 140 GCTGGCCGGCTGCATCGCGGCCACCGGGATGACATGTGGAGCAGCCAGGACCTGTACGA 199
 Db 61 GCTGGCCGGCTGCATCGCGGCCACCGGGATGACATGTGGAGCAGCCAGGACCTGTACGA 120
 QY 200 CAAACCCGTCACCTCGGTGTTCCAGTACGAAGGGCTCTGGAGAGCTGGTGGAGCAGAG 259
 Db 121 CAAACCCGTCACCTCGGTGTTCCAGTACGAAGGGCTCTGGAGAGCTGGTGGAGCAGAG 180
 QY 260 TTCAGGCTTCACGATGAGCCCTTATTTACCATCTCTGGAGCTTCAGCCATGCTGCA 319
 Db 181 TTCAGGCTTCACGATGAGCCCTTATTTACCATCTCTGGAGCTTCAGCCATGCTGCA 240
 QY 320 GGCAGTGGAGCCCTGATGATCGTAGGATCGTCTGGTGGTCCATGGCCTCTGTATC 379
 Db 241 GGCAGTGGAGCCCTGATGATCGTAGGATCGTCTGGTGGTCCATGGCCTCTGTATC 300
 QY 380 CATCTTTCCCTGAAATGCATCCGATTCGGAGCATGGAGGACTCTGCCAAAGCCAAAT 439
 Db 301 CATCTTTCCCTGAAATGCATCCGATTCGGAGCATGGAGGACTCTGCCAAAGCCAAAT 360
 QY 440 GACACTGACCTCCGGATCATGTTCAATGTCAGGTCTTTGTGCAATTTGTGGAGTCTC 499
 Db 361 GACACTGACCTCCGGATCATGTTCAATGTCAGGTCTTTGTGCAATTTGTGGAGTCTC 392
 QY 500 TGTGTTTCCCAACATGCTGGTGAACCTTCTGGATGTCACAGCTAACATGTACACCGG 559

TITLE

JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 TITLE

1383	CCCTAGATGAATGAGAAAAATTAATTTTTTTAAATTTAAAGTCTCTAAATATATAGTTAAATAA	1444
1640	TAATGTTTTAGTAAAAATGATACACTATCTCTGTGAAATAGCCCTCAACCCCTACATGTGGAT	1699
1443	TAATGTTTTAGTAAAAATGATACACTATCTCTGTGAAATAGCCCTCACCCCTACATGTGGAT	1502
1700	AGAAGGAATGAAAAAATAATTCGTTTGACATGTTCTATATGTGTACTTTTGTAAAGTCATG	1759
1503	AGAAGGAATGAAAAAATAATTCGTTTGACATGTTCTATATGTGTACTTTTGTAAAGTCATG	1562
1760	CTTAAGTACAAATTCATGAAAA-----	1782
1563	CTTAAGTACAAATTCATGAAAAAGCTCACTGATCCTTAATTTCTTCCCTTTGAGGTCTCTA	1622
1783	-----	1782
1623	TGGCTCTGATTTGATACATAGTAGTAAGTGAAGCCATGTAAAAAGTAAATAATGTCTGGC	1682
1783	-----GCTCACACCTGTAATTCCTAGCACTTTGGGAGGCTGAGGAGGAGGATCAGTGA	1836
1683	ACAGTGCTCTACGCTCTGTAATTCCTAGCACTTTTCGGAGGCTGAGGAGGAGGATCGCTGA	1742
1837	GCCCAGAGTTTCGAGACTAGCCTCGGCGAAATGAGAGAGCCCTCTCTCTACAAAATACAG	1896
1743	GCCCAGAGTTTCGAGACTAGCCTCGGCGAAATGAGAGAGCCCTCTCTCTACAAAATACAG	1802
1897	AGAAAAAATCAGCCAGTTCATGTGGCGATACCTGTAGTCCCGAGTCCCGGAGGCT	1956
1803	AGAAAAAATCAGCCGCTCATGTGCGATACACCTGTAGTCCCGAGTCCCGGAGGCT	1862
1957	GAGTGGGAGGATCACTTGAGCCAGGAGGTTGGGCTGAGTGGCCATGATCACACC	2016
1863	GAGTGGGAGGATCGCTTGAGCCAGGAGGTTGGGCTGAGTGGCCATGATCACACC	1922
2017	ACTGCATCTCCAGCCAGGTGACATAGCAGATCCCTGTCTAAAAAATAAAAAATAAATAAT	2076
1923	ACTGCATCTCCAGCCAGGTGACATAGCAGATCCCTGTCTAAAAAATAAAAAATAAATAAT	1987
2077	GGAACACAGCAAGTCTAGGAAGTAGGTTAAAACTAAATTCCTTAA	2121
1983	GGAACACAGCAAGTCTAGGAAGTAGGTTAAAACTAAATTCCTTAA	2027
RESULT 8		
AC016252/c		
LOCUS	164953 bp DNA linear	PRI 25-SEP-2000
DEFINITION	Homo sapiens 3 BAC RP11-461M2 (Roswell Park Cancer Institute Human BAC Library) complete sequence.	
ACCESSION	AC016252	
VERSION	AC016252.21	GI:19774335
KEYWORDS	HTG.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	1 (bases 1 to 164953) Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-ouman, F.R., Allen, C., Alshrook, S.L., Amarantunge, H.C., Are, J.R., Ayelle, M., Banks, T., Barthelia, J., Benton, J., Biname, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., He, X., Hernandez, J., Hernandez, O., Hodgson, A., Hoques, M., Holloway, C.,	

Hollins, B., Honsai, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Louibeged, H., Lozado, R., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Moore, S., Morgan, M., Moorish, T., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I., Sodergren, E., Sonaika, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Zorrilla, S., Naylor, S.L., Weinstein, G., and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 164953)
Worley, K.C.
Direct Submission
Submitted (24-NOV-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 164953)
Worley, K.C.
Direct Submission
Submitted (06-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 164953)
Worley, K.C.
Direct Submission
Submitted (28-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 164953)
Worley, K.C.
Direct Submission
Submitted (25-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
6 (bases 1 to 164953)
Worley, K.C.
Direct Submission
Submitted (25-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Mar 28, 2002 this sequence version replaced gi:19172796.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES	source	Location/Qualifiers
		1..164953
		/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"
		/chromosome="3"
		/clone="RP11-461M2"
		/note="overlaps bases 1..2055 of clone AC023049"
		/function="clone overlap"
		/complement(81..571)
		/rpt_family="L1MC/D"
		/complement(802..1221)
		/rpt_family="L1M4"
		/complement(1351..3211)
		/rpt_family="L1MB2"
		/complement(4306..4615)
		/rpt_family="AluSx"
		5734..5828
		/rpt_family="L2"
		5932..6109
		/rpt_family="L2"
		/complement(6172..6471)
		/rpt_family="AluJb"
		6496..7131
		/rpt_family="L1MC/D"
		6500..7300
		/function="trinucleotide tandem repeat"
		7144..7401
		/rpt_family="L1MC/D"
		7509..7591
		/rpt_family="L1MC/D"
		/complement(9509..9612)
		/rpt_family="L1MB6"
		/complement(10594..10729)
		/rpt_family="L1MC/D"
		10810..11005
		/rpt_family="L1MC/D"
		11077..11124
		/rpt_family="L2"
		11922..11943
		/rpt_family="AT-rich"
		12284..12589
		/rpt_family="AluSx"
		12914..13614
		/rpt_family="L1MC/D"
		/complement(14500..14605)
		/rpt_family="HAL1"
		/complement(16916..17151)
		/rpt_family="L1MB6"
		/complement(17156..18175)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 182726)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Cattle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihoval,T., Menga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (04-NOV-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 182726)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihoval,T., Menga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jan 15, 2001 this sequence version replaced gi:11094829.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RW/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www.seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L11425
 Center clone name: 468_E 18
 ----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 168531 bases at least Q40
 Consensus quality: 175180 bases at least Q30

Consensus quality: 177830 bases at least Q20
 Insert size: 179326; sum-of-contigs
 Quality coverage: 4.5 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently consists of 35 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
 * 1 1676: contig of 1676 bp in length
 * 1677: gap of 100 bp
 * 1777: contig of 891 bp in length
 * 2668: gap of 100 bp
 * 2768: contig of 353 bp in length
 * 3121: gap of 100 bp
 * 3221: contig of 699 bp in length
 * 3919: gap of 100 bp
 * 4019: gap of 100 bp
 * 4921: contig of 902 bp in length
 * 5021: gap of 100 bp
 * 6412: contig of 1391 bp in length
 * 6512: gap of 100 bp
 * 7808: contig of 1296 bp in length
 * 7908: gap of 100 bp
 * 9082: contig of 1174 bp in length
 * 9182: gap of 100 bp
 * 9183: contig of 2336 bp in length
 * 11518: gap of 100 bp
 * 11519: contig of 2809 bp in length
 * 14228: gap of 100 bp
 * 14528: contig of 1913 bp in length
 * 16441: gap of 100 bp
 * 16541: contig of 2800 bp in length
 * 19341: gap of 100 bp
 * 19441: contig of 3166 bp in length
 * 22607: gap of 100 bp
 * 22707: contig of 2509 bp in length
 * 25216: gap of 100 bp
 * 25316: contig of 2535 bp in length
 * 27851: gap of 100 bp
 * 27951: contig of 19970 bp in length
 * 47921: gap of 100 bp
 * 48021: contig of 2592 bp in length
 * 50613: gap of 100 bp
 * 50713: contig of 2345 bp in length
 * 53058: gap of 100 bp
 * 53158: contig of 1703 bp in length
 * 54861: gap of 100 bp
 * 54961: contig of 2085 bp in length
 * 57046: gap of 100 bp
 * 57146: contig of 3830 bp in length
 * 60976: gap of 100 bp
 * 61076: contig of 4857 bp in length
 * 65933: gap of 100 bp
 * 66033: contig of 4012 bp in length
 * 70044: gap of 100 bp
 * 70145: contig of 6437 bp in length
 * 76582: gap of 100 bp
 * 76682: contig of 6806 bp in length
 * 83488: gap of 100 bp
 * 83588: contig of 5447 bp in length
 * 89035: gap of 100 bp
 * 89135: contig of 6276 bp in length
 * 95110: gap of 100 bp
 * 95511: contig of 8055 bp in length
 * 103566: gap of 100 bp
 * 103666: contig of 8930 bp in length
 * 112596: gap of 100 bp
 * 112696: contig of 9965 bp in length
 * 122661: gap of 100 bp
 * 122761: contig of 8991 bp in length


```

* 131752 131851: gap of 100 bp
* 131852 144507: contig of 12656 bp in length
* 144508 144607: gap of 100 bp
* 144608 160481: contig of 15874 bp in length
* 160482 160581: gap of 100 bp
* 160582 181386: contig of 20805 bp in length
* 181387 181486: gap of 100 bp
* 181487 182726: contig of 1240 bp in length.

```

FEATURES

```

source
1..182726
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/map="3"
/clone="RP11-468E18"
/clone_lib="RPC1-11 Human Male BAC"
1..1676
/note="assembly_fragment"
clone_end:SP6
vector_side:left
1777..2667
/note="assembly_fragment"
2768..3120
/note="assembly_fragment"
3221..3919
/note="assembly_fragment"
4020..4921
/note="assembly_fragment"
5022..6412
/note="assembly_fragment"
6513..7808
/note="assembly_fragment"
7909..9082
/note="assembly_fragment"
9183..11518
/note="assembly_fragment"
11619..14427
/note="assembly_fragment"
14528..16440
/note="assembly_fragment"
16541..19340
/note="assembly_fragment"
19441..22606
/note="assembly_fragment"
22707..25215

```

```

Query Match      61.4%; Score 1303.2; DB 2; Length 182726;
Best Local Similarity 93.0%; Pred. No. 6.6e-287;
Matches 1418; Conservative 0; Mismatches 3; Indels 103; Gaps 1;

QY 701 CAACCTACAAAGCGGTTCTTATCATGCTCAGCCACAGTGTGCTCAAGCCTGGAGG 760
DB 47532 CAGCTACAAAGCGGTTCTTATCATGCTCAGCCACAGTGTGCTCAAGCCTGGAGG 47473

QY 761 CTTCAAGGCCAGCAGCTGGCTTTGGGTCCAAACACCAAAACAAAGAGATATACGATGGAGG 820
DB 47472 CTTCAAGGCCAGCAGCTGGCTTTGGGTCCAAACACCAAAACAAAGAGATATACGATGGAGG 47413

QY 821 TCCCGCCACAGAGGAGGATCAATCTTATCTTCCAGCAGCAGTATGTGTATGCTC 880
DB 47412 TCCCGCCACAGAGGAGGATCAATCTTATCTTCCAGCAGCAGTATGTGTATGCTC 47353

QY 881 TAAAGACTCTCAGCAGCGGGAAGAACTCCCGGAGAGCTACCCCAAAAAACAAAGGAGA 940
DB 47352 TAAAGACTCTCAGCAGCGGGAAGAACTCCCGGAGAGCTACCCCAAAAAACAAAGGAGA 47293

QY 941 TCCCATCTAGATTTCTTCTTGTCTTTTGAAGTCAAGCTGGAAGTTAGAAAAGCCTCGATTT 1000
DB 47292 TCCCATCTAGATTTCTTCTTGTCTTTTGAAGTCAAGCTGGAAGTTAGAAAAGCCTCGATTT 47233

QY 1001 CATCTTTGGAGAGCCAAATGGTCTTAGCTCAGTCTCTGTCTCTAAATATCCACATA 1060

```

```

Db 47232 CATCTTTGGAGAGCCAAATGGTCTTAGCTCAGTCTCTGTCTCTAAATATCCACATA 47173
QY 1061 AAACAGCTGAGTTATTTATGAATTTAGAGGCTATAGCTACAAATTTTCAATCTCTATTCT 1120
DB 47172 AAACAGCTGAGTTATTTATGAATTTAGAGGCTATAGCTACAAATTTTCAATCTCTATTCT 47113
QY 1121 TTTTAAATATAAATCTTCTCTGATGAGAGATGGTGTAACTCTCTCTCATCAT 1180
DB 47112 TTTTAAATATAAATCTTCTCTGATGAGAGATGGTGTAACTCTCTCTCATCAT 47053
QY 1181 TTTGATGATTTAGACAGACTCCCTCTCTCTCTCTAGTCAATAAACCATTGATGATCTA 1240
DB 47052 TTTGATGATTTAGACAGACTCCCTCTCTCTCTAGTCAATAAACCATTGATGATCTA 46993
QY 1241 TTTCCAGCTTATCCCAAGAAAATTTTGAAGAAAAGAGTAGACCCAAAGATGTTATT 1300
DB 46992 TTTCCAGCTTATCCCAAGAAAATTTTGAAGAAAAGAGTAGACCCAAAGATGTTATT 46933
QY 1301 TTTCTGCTGTTGAATTTTGTCTCCCAACCCCTTGTAGTAAATAAACAACATCTACTGA 1360
DB 46932 TTTCTGCTGTTGAATTTTGTCTCCCAACCCCTTGTAGTAAATAAACAACATCTACTGA 46873
QY 1361 AGAAGAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGTTTTCTT 1420
DB 46872 AGAAGAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGTTTTCTT 46813
QY 1421 ACACCTGTGATCTTAAAGTTACCAAAACCAAAAGTCAATTTTCTAGTTTGGAGCAACCAACCT 1480
DB 46812 ACACCTGTGATCTTAAAGTTACCAAAACCAAAAGTCAATTTTCTAGTTTGGAGCAACCAACCT 46753
QY 1481 TTCTACTGCTGTTGACATCTTCTTATACAGCAACCAACATCTAGAGTTTCTCAGCTC 1540
DB 46752 TTCTACTGCTGTTGACATCTTCTTATACAGCAACCAACATCTAGAGTTTCTCAGCTC 46693
QY 1541 TCCACTGGAGTCTCTTTCTGTGGGGTTCAGAAATTCCTCCTAGATGAATGAGAAAAT 1600
DB 46692 TCCACTGGAGTCTCTTTCTGTGGGGTTCAGAAATTCCTCCTAGATGAATGAGAAAAT 46633
QY 1601 ATTTTAAATTTAAAGTCTTAAATATATAGTTTAAATAAATAATGTTTGTAGTAAATGATA 1660
DB 46632 ATTTTAAATTTAAAGTCTTAAATATATAGTTTAAATAAATAATGTTTGTAGTAAATGATA 46573
QY 1661 CACTATCTCTGTAATAGCTCCACCCCTACATGTTGGATAGAGAAATGAAAAAATAAT 1720
DB 46572 CACTATCTCTGTAATAGCTCCACCCCTACATGTTGGATAGAGAAATGAAAAAATAAT 46513
QY 1721 TGCTTTGACATTTCTATATGTTTCTTGTAAAGTCAATGTTAAGTACAAATTCATGAA 1780
DB 46512 TGCTTTGACATTTCTATATGTTTCTTGTAAAGTCAATGTTAAGTACAAATTCATGAA 46453
QY 1781 AA----- 1782
DB 46452 AAGCTCACTGATCTAAATTTCTTCCCTTTGAGGTCTCTATGGCTCTGATTTGATGATA 46393
QY 1783 -----GCTCACACCTCTGTAAT 1797
DB 46392 GTAAGTGTAAAGCAATGTAAAAAGTAAATAATGTCTGGGCACAGTGCTCAGCCTGTAAT 46333
QY 1798 CTTAGCACCTTTGGGAGGCTGAGGAGGAGGATCACTTTGAGCCCGAGAGTTTCGAGACTAGC 1857
DB 46332 CTTAGCACCTTTGGGAGGCTGAGGAGGAGGATCACTTTGAGCCCGAGAGTTTCGAGACTAGC 46273
QY 1858 CTGGGCAACATCGAGAGAGCCCTGTCTCTAATAAATAAGAGAAATAATCAGCCAGTCA 1917
DB 46272 CTGGGCAACATCGAGAGAGCCCTGTCTCTAATAAATAAGAGAAATAATCAGCCAGTCA 46213
QY 1918 TGGTGGCATACACCTGTAGTCCCGAGGCTGAGTGGGAGGATCACTTTGAG 1977
DB 46212 TGGTGGCCCTACACCTGTAGTCCCGAGGCTGAGTGGGAGGATCACTTTGAG 46153
QY 1978 CCCAGGAGGAGTGGGGCTGAGTGGAGCCATGATCACACCACTGCATCTCCAGCCAGGTGAC 2037
DB 46152 CCCAGGAGGAGTGGGGCTGAGTGGAGCCATGATCACACCACTGCATCTCCAGCCAGGTGAC 46093

```



```

misc_feature 46985..55777
              /note="assembly_fragment"
misc_feature 55878..63178
              /note="assembly_fragment"
misc_feature 63279..73509
              /note="assembly_fragment"
misc_feature 73610..82711
              /note="assembly_fragment"
misc_feature 82812..93642
              /note="assembly_fragment"
misc_feature 93743..109313
              /note="assembly_fragment"
misc_feature 109414..124437
              /note="assembly_fragment"
misc_feature 124538..145477
              /note="assembly_fragment"

ORIGIN
Query Match      51.1%; Score 1083.2; DB 2; Length 145477;
Best Local Similarity 99.7%; Pred. No. 1.4e-236;
Matches 1085; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 701 CAACACAAAGCGGTTCTTATCATGCTCAGGCCACAGTGTGCTCAAGCCCTGGAGG 760
Db 18143 CAGCTACAAAGCGGTTCTTATCATGCTCAGGCCACAGTGTGCTCAAGCCCTGGAGG 18202

Qy 761 CTTCAAGCCAGCAGCTGGCTTGGCTCCAAACCAACCAAGAGATATAGATGGAGG 820
Db 18203 CTTCAAGCCAGCAGCTGGCTTGGCTCCAAACCAACCAAGAGATATAGATGGAGG 18262

Qy 821 TGCCCGCAGACGAGGAGTACAAATCTTATCTTCCAGCAGCAGTATGTGTAATGCTC 880
Db 18263 TGCCCGCAGACGAGGAGTACAAATCTTATCTTCCAGCAGCAGTATGTGTAATGCTC 18322

Qy 881 TAAGACCTCTCAGCAGCGGCGAAGAACTCCCGAGAGCTCACCCAAAACAAAGGAGA 940
Db 18323 TAAGACCTCTCAGCAGCGGCGAAGAACTCCCGAGAGCTCACCCAAAACAAAGGAGA 18382

Qy 941 TCCCATCTAGATTTCTTCTGCTTTTGGCTCAGCTGGAAGTGAAGAGCTCGATT 1000
Db 18383 TCCCATCTAGATTTCTTCTGCTTTTGGCTCAGCTGGAAGTGAAGAGCTCGATT 18442

Qy 1001 CATCTTTGAGAGGCGCAATGTCTTAGCTCAGTCTCTGTCTCTAAATATTCACCATTA 1060
Db 18443 CATCTTTGAGAGGCGCAATGTCTTAGCTCAGTCTCTGTCTCTAAATATTCACCATTA 18502

Qy 1061 AAACAGCTGAGTTATTTAAGATTAAGAGCTATAGCTGACATTTCAATCTCTATTCT 1120
Db 18503 AAACAGCTGAGTTATTTAAGATTAAGAGCTATAGCTGACATTTCAATCTCTATTCT 18562

Qy 1121 TTTTAAATATAAATCTTCTACTCTGATGAGAGATGTGTTTAACTCTCTCTCACAT 1180
Db 18563 TTTTAAATATAAATCTTCTACTCTGATGAGAGATGTGTTTAACTCTCTCTCACAT 18622

Qy 1181 TTTGATGATTTAGACAGACTCCCTCTCTCTAGTCAATTAACCCATGATGATCTA 1240
Db 18623 TTTGATGATTTAGACAGACTCCCTCTCTCTAGTCAATTAACCCATGATGATCTA 18682

Qy 1241 TTTCCAGCTTATCCCAAGAAACTTTTGAAGAGAGAGTAGAGCCCAAGATGTTAT 1300
Db 18683 TTTCCAGCTTATCCCAAGAAACTTTTGAAGAGAGAGTAGAGCCCAAGATGTTAT 18742

Qy 1301 TTTCTGCTGTTGAATTTGTTCTCCCAACCCCACTTGGCTAGTAAATACACTTACTGA 1360
Db 18743 TTTCTGCTGTTGAATTTGTTCTCCCAACCCCACTTGGCTAGTAAATACACTTACTGA 18802

Qy 1361 AGAAGAGCAATTAAGAGAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTCTT 1420
Db 18803 AGAAGAGCAATTAAGAGAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTCTT 18862

Qy 1421 ACAGTGTGATCTTAAAGTTTACCAACCAAGTCAATTTTCAGTTTCAGGCAACCAACCT 1480
Db 18863 ACAGTGTGATCTTAAAGTTTACCAACCAAGTCAATTTTCAGTTTCAGGCAACCAACCT 18922

```

```

Qy 1481 TTCTACTGCTGTTGACATCTTTATTATACAGCAACACATTTCTAGGAGTTTCTTGAGCTC 1540
Db 18923 TTCTACTGCTGTTGACATCTTTATTATACAGCAACACATTTCTAGGAGTTTCTTGAGCTC 18982

Qy 1541 TCACCTGGAGTCTCTTTCTGTCGCGGGTCAGAAATTTGTCCTAGATCAATGAGAAAAATT 1600
Db 18983 TCACCTGGAGTCTCTTTCTGTCGCGGGTCAGAAATTTGTCCTAGATCAATGAGAAAAATT 19042

Qy 1601 ATTTTATTTTAAATTAAGTCTTAAATATAGTTTAAATAAATATGTTTGTAGTAAATGATA 1660
Db 19043 ATTTTATTTTAAATTAAGTCTTAAATATAGTTTAAATAAATATGTTTGTAGTAAATGATA 19102

Qy 1661 CACTATCTCTGTGAAATAGCCTCACCCCTACATGTCGATAGAGAAATGAAAAAATAT 1720
Db 19103 CACTGTCTCTGTGAAATAGCCTCACCCCTACATGTCGATAGAGAAATGAAAAAATAT 19162

Qy 1721 TGCTTTGACATGTTCTATATGTTTGTAAAGTTCATGCTTAAAGTTCATGCTTAAAGTTCATGCT 1780
Db 19163 TGCTTTGACATGTTCTATATGTTTGTAAAGTTCATGCTTAAAGTTCATGCTTAAAGTTCATGCT 19222

Qy 1781 AAGCTCAC 1788
Db 19223 AAGCTCAC 19230

RESULT 11
BD237992 982 bp DNA linear PAT 17-JUL-2003
LOCUS Gastric polypeptide ZSIG28.
DEFINITION BD237992
ACCESSION BD237992.1 GI:33047762
VERSION JP 2002524103-A/1.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 982)
AUTHORS Sheppard, P.O. and Foley, K.P.
TITLE Gastric polypeptide ZSIG28
JOURNAL Patent: JP 2002524103-A 1 06-AUG-2002;
        ZYMOGENETICS INC
COMMENT OS Homo sapiens (human)
        PN JP 2002524103-A/1
        PD 06-AUG-2002
        PF 14-SBP-1999 JP 2000570197
        PR 16-SBP-1998 US 09/154444
        PI PAUL O SHEPPARD, KEVIN P FOLEY
        PC C12N15/09,A61K38/00,A61K39/395,A61K39/395,A61P1/04,A61P1/14,
        PC A61P3/08,
        PC A61P5/50,A61P31/04,A61P31/10,A61P35/00,C07K14/47,C07K16/18, PC
        C12N1/15,
        PC C12N1/19,C12N1/21,C12N5/10,C12P21/02,C12P21/08,C12Q1/02 CC
        C12Q1/68,G01N33/15,
        PC G01N33/50,G01N33/53,G01N33/577,C12N15/00,C12N5/00,A61K37/02 CC
        Gastric polypeptide ZSIG28
        FH Key Location/Qualifiers
        FT CDS
            (70)..(853)
FEATURES
    source
        1..982
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
ORIGIN
Query Match      45.0%; Score 954.2; DB 6; Length 982;
Best Local Similarity 99.1%; Pred. No. 1.6e-207;
Matches 970; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

Qy 21 TTAGCTTTCACACTTCGCGAGCAGGAGCGGCGAGCTTCTCGAGCGGCGGCGGCG 80
Db 1 TTAGCTTTCACACTTCGCGAGCAGGAGCGGCGAGCTTCTCGAGCGGCGGCGGCG 60

```


Qy	1153	GAATGGGTTTAAATCTCTCTCACATTTTGTGATGATTAGACAGACTCCCCCTCTTCCT	1212
Db	1153	GAATGGAAATTTGACTCACAGACTAATACTTTTAATGGTTTAGAGAACTTTC--CTTCCT	1209
Qy	1213	CCTAGTCAATAAACCCATTGGATGATCTATTTCCAGCTT-ATCCCAAGAAAACTTTTGA	1271
Db	1210	CGTACTTAATGAAGCCTGCTGAATGTCGATTTTCAGCTTGACCACCAAGGGAATTTTAA	1269
Qy	1272	AAGGAAGAGTAGACCCAAAGATGTTATTTTCTGCTGTTGAAATTTTGTCTCCCAACCC	1331
Db	1270	AAGGAAAAAATAATACATTTAAAGGCAATATTTCCTACTCAATGTGCCTTACCCACCC	1329
Qy	1332	CAACTTGGCTAGTAATA-----AACACTTACTGAAGAAGCAATPAGAGAAAGATAT	1385
Db	1330	CAACTTGACTGATATAATAATGAACACCACTTAAAGAAAAAATGCCAGAGGAAAGATAG	1389
Qy	1386	TTGT-AAATCTCTCCAGGCCCATGATCTCGGTTTTCTTACACTGTGATCTTAAAAGTTTACCA	1444
Db	1390	TTGTGTTTTCCCCAGCCAGTCATCTGAGTCCCGCTATGTGGTGATCTAGAACATTA-CT	1448
Qy	1445	AACCAAGTCAATTTTCAGTTTGAGGCCAACCAACCTTTCTACTGCTGTTGACATCTTCTT	1504
Db	1449	CGCCACAGTGAATTTTCA-AAGAAAGGCAAGCGAGCCTGTTGGCTCTGCTCAGCATCTGCTG	1507
Qy	1505	ATTACAGCAACACCATTTAGGAGTTTCTTGAGCTCTCCACTGGAGTCTCTTTCTGTGCG	1564
Db	1508	ATTCAGCAGGCCCCTTCAGAGCTTTCCA-----CTAGAAGTCTCTCTTCTCTCG	1558
Qy	1565	CGGGTCAGAAATTTGCTCCTAGATGAATGAGAAA---ATTATTTTTTTTAAATTTAAGTCCT	1621
Db	1559	GAAGTCAGAAATTTCCCCCTAGAAGAGTAGAATAAGATCTTTTGGGTAACTGAGTCTT	1618
Qy	1622	AAATATAGTTTAAATAAATAATGTTTTAGTAAATGATACATATCTCTGTGAAATAGCC	1681
Db	1619	AGGTATAGTTTATATAAATAAGTATATATAGCAAAACGGTTTGGTATCTCAGTGAATTAGTT	1678
Qy	1682	TCACCCCTCATGTTGGATAGA-AGGAAATGAAAAATAATTGCTTTTGACATTTGCTATAT	1740
Db	1679	TCAGCCTTACATATAGAAAAAGCTGGGGAAAAAANAAGCATCCCTTGACATTTGCTATAG	1738
Qy	1741	GGTACTTTTGAAGTCATGCTTAAAGTACAA	1770
Db	1739	CGTAAAGATCCTATATAAATCCCAAGCTTCAA	1768

```

RESULT 14
AF221069
LOCUS      AF221069
DEFINITION Homo sapiens Claudin-18 mRNA, complete cds.
ACCESSION AF221069
VERSION   AF221069.1
KEYWORDS  GI:6715517
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS   Niimi,T., Nagashima,K., Ward,J.M., Minoo,P., Zimonjic,D.B.,
            Popescu,N.C. and Kimura,S.
TITLE      claudin-18, a novel downstream target gene for the T/EBP/NKX2.1
            homeodomain transcription factor, encodes lung- and
            stomach-specific isoforms through alternative splicing
JOURNAL   Mol. Cell. Biol. 21 (21), 7380-7390 (2001)
MEDLINE   21470315
PUBMED    11585919
REFERENCE 2 (bases 1 to 816)
AUTHORS   Niimi,T. and Kimura,S.
TITLE      Direct Submission
JOURNAL   Submitted (04-JAN-2000) Laboratory of Metabolism, DBS, National
            Cancer Institute, NIH, 9000 Rockville Pike, Bethesda, MD 20892, USA
FEATURES  Location/Qualifiers
            1..816
            /organism="Homo sapiens"
source

```

```

CDS
/mol_type="mRNA"
/db_xref="taxon:9606"
10..795
/codon_start=1
/product="Claudin-18"
/protein_id="AAP26448.1"
/db_xref="GI:6715518"
/translation="MSTTTCQVAFLLSLIGLACTAATGCMQMTQDLYDNPTVTSF
QYGLNRSCVQSSGTECRPTTILGLPAMQAVRALMIVGILGAQLGLVSIIPALK
CIRIGMEDSAKANMTTIGMFIIVSGLCAIAGVSFANMLVTFMSTANNYTGMGG
MVQTVQTRYTFGAALFVGVAGGLTLIGVVMCMCIACRGLAPEPTNYKAVSYHASGHSV
AYRPGFSKATSTQNKXKIKYDGGARTEDEVQSPSKHDVVV"

```

ORIGIN

Query Match	38.5%	Score 816	DB 9	Length 816
Best Local Similarity	100.0%	Prod. No. 6.5e-176		
Matches 816	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	81	GCACGAGTATGTCACACACCATGCGGAGTGGTGGCGTTCCTCTGTCTCATCTCTGGGG	140	
Db	1	GCACGAGTATGTCACACACCATGCGGAGTGGTGGCGTTCCTCTGTCTCATCTCTGGGG	60	
QY	141	CTGGCCGGCTGCATCGCGGCACCGGAGTGGACATGTGGAGCACCCAGACACTGTACGAC	200	
Db	61	CTGGCCGGCTGCATCGCGGCACCGGAGTGGACATGTGGAGCACCCAGACACTGTACGAC	120	
QY	201	AACCCCGTCACCTCCGTGTTCAGTAGCAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGT	260	
Db	121	AACCCCGTCACCTCCGTGTTCAGTAGCAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGT	180	
QY	261	TCAGGCTTCACGAATGCGAGGCCCTAATTTCAATCCTCGGACTTCCAGCCATGCTGCAG	320	
Db	181	TCAGGCTTCACGAATGCGAGGCCCTAATTTCAATCCTCGGACTTCCAGCCATGCTGCAG	240	
QY	321	GCACTGCGAGCCCTGATGATCTGAGGATCGTCTGGGTGCCATTTGGGCTCTCTGTATCC	380	
Db	241	GCACTGCGAGCCCTGATGATCTGAGGATCGTCTGGGTGCCATTTGGGCTCTCTGTATCC	300	
QY	381	ATCTTTGCCCTCAAAATGCAATCGGCATTTGGCAGCATGGAGGACTCTGCCAAAGCCAAACATG	440	
Db	301	ATCTTTGCCCTCAAAATGCAATCGGCATTTGGCAGCATGGAGGACTCTGCCAAAGCCAAACATG	360	
QY	441	ACACTGACCTCGGGATCATGTTCATTTGTCTCAGTCTTTGTGCAATTCCTGGAGTGTCT	500	
Db	361	ACACTGACCTCGGGATCATGTTCATTTGTCTCAGTCTTTGTGCAATTCCTGGAGTGTCT	420	
QY	501	GTGTTTGCCAAATGCTGGTGACTAACTCTGGATGTCCAGCTAAATGTACACCGGC	560	
Db	421	GTGTTTGCCAAATGCTGGTGACTAACTCTGGATGTCCAGCTAAATGTACACCGGC	480	
QY	561	ATGGGTGGGATGGTGACAGCTGTTCCAGACCAAGGTACACATTTGGTGGCGCTCTGTCTGTG	620	
Db	481	ATGGGTGGGATGGTGACAGCTGTTCCAGACCAAGGTACACATTTGGTGGCGCTCTGTCTGTG	540	
QY	621	GGCTGGGTGCTGGAGGCGTCACTAATTTGGGGGTGTGATGTGATGTCATCGCTGCCGG	680	
Db	541	GGCTGGGTGCTGGAGGCGTCACTAATTTGGGGGTGTGATGTGATGTCATCGCTGCCGG	600	
QY	681	GGCTGGGACAGAGAAACCACTACAAAGCGTTTCTTATCATGTCTCAGGCCACAGT	740	
Db	601	GGCTGGGACAGAGAAACCACTACAAAGCGTTTCTTATCATGTCTCAGGCCACAGT	660	
QY	741	GTTGCTTCAAGCCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCACACCAAAAC	800	
Db	661	GTTGCTTCAAGCCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCACACCAAAAC	720	
QY	801	AAGAAGATATACGATGGAGGTGCCCGCACAGGACGAGGTACAATCTTATCTTCCAG	860	
Db	721	AAGAAGATATACGATGGAGGTGCCCGCACAGGACGAGGTACAATCTTATCTTCCAG	780	
QY	861	CACGACTATGTGTAATGCTTAAGACCTCTCAGCAC	896	
Db	781	CACGACTATGTGTAATGCTTAAGACCTCTCAGCAC	816	

RESULT 15
AF349450
LOCUS
DEFINITION
Mus musculus claudin-18A1.2 mRNA, complete cds, alternatively
spliced.
ACCESSION
AF349450
VERSION
AF349450.1 GI:16224162
KEYWORDS
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1409)
REFERENCE
AUTHORS
Niimi,T., Nagashima,K., Ward,J.M., Minoo,P., Zimonjic,D.B.,
Popescu,N.C. and Kimura,S.
TITLE
claudin-18, a novel downstream target gene for the T/EBP/NKX2.1
homeodomain transcription factor, encodes lung- and
stomach-specific isoforms through alternative splicing
Mol. Cell. Biol. 21 (21), 7380-7390 (2001)
JOURNAL
MEDLINE
PUBMED
11585919
REFERENCE
AUTHORS
Niimi,T. and Kimura,S.
TITLE
Direct Submision
JOURNAL
Submitted (15-FEB-2001) Laboratory of Metabolism, National Cancer
Institute, NIH, 9000 Rockville Pike, Bethesda, MD 20892, USA
FEATURES
source
1..1409
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/note="isolated from mixed background of 129/Sv, C57BL/6,
and Black Swiss mice"
62..688
/note="alternatively spliced"
/codon_start=1
/product="claudin-18A1.2"
/protein_id="AA15635.1"
/db_xref="GI:16224163"
/translation="WATTTTCQVVGLLSLGLAGCIATGMDMSTQDLYNPVTAVF
QYGLWRSCQSSGFTCEPFTTILGLPAMLOAVRALMIVGLVIGILVSIIFALK
CIRIGMSDASAKMTLFIISGICAIIGSVFANMLVTFNFMSTANMYSQWGG
MGGVQTVQTRTFGAALFVGVAGLTLIGVWMCIAICRGLTPDDSK"
ORIGIN
Query Match 34.6%; Score 733.4; DB 10; Length 1409;
Best Local Similarity 76.1%; Pred. No. 5.8e-157;
Matches 1047; Conservative 0; Mismatches 286; Indels 42; Gaps 10;
QY 36 CGGACGAGGCGGAGCTTCTCGAGGCGGAGGCGGCGGCGGAGGATCATGTCC 95
DB 10 CAGCAAGAGGTTGGTGTCTCTCGAGAGCTGC--GCCACGAGCTGAGCCATGGCC 67
QY 96 ACCACACATGCCAAGTGTGGCGTTCTCTCTGTCATCTGCGGCGGCGGCGGCGGCTGCATC 155
DB 68 ACCACACATGCCAAGTGTGGCGTTCTCTCTGTCATCTGCGGCGGCGGCGGCTGCATC 127
QY 156 GCGGCGACCGGAGTGCATGTGGAGCACCAGGACCTGTACACAAACCCGCTCAGCTCC 215
DB 128 GCGGCGACCTGGATGACATGTGGAGCACTCAAGACCTGTATGACAAACCCAGTCAACGCC 187
QY 216 GTCTTCCAGTACAGGCGCTCTGAGGAGCTGGTGGAGGAGGATTCAGGCTTCACCGAA 275
DB 188 GTGTTCAGTATGAGGGCTCTGAGAGGTTGGTGCAACAGAGCTGGGGTTTACCGAG 247
QY 276 TGCAGGCGCTATTTACACATCTCGGAGCTTCCAGCCATGCTGAGGAGCTGGAGCGCTG 335
DB 248 TGCAGGCGCTATTTACACATCTCGGAGCTTCCAGCCATGCTGCAAGCTGTACGAGCGCTG 307
QY 336 ATGATCTGAGGATCTGCTGGGTGCGATTTGGCTTCTGATATCCATCTTTGGCTGAAA 395

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2004, 14:08:53 ; Search time 118 Seconds
(without alignments)
9975.007 Million cell updates/sec

Title: US-10-063-731-117
Perfect score: 2121
Sequence: 1 gagctccctcaggagcgcg.....gggttaaaactaattctttaa 2121

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	733	34.6	1353	4	US-09-724-864-32
2	174.4	8.2	3982	4	US-09-820-924-3
3	170.8	8.1	9365	4	US-09-608-285A-8
4	170.8	8.1	9365	4	US-09-350-836B-8
5	170.8	8.1	9365	4	US-09-370-265-8
6	170.8	8.1	9365	4	US-09-557-800C-8
7	170.8	8.1	9365	4	US-09-370-625A-8
8	170.8	8.1	14747	4	US-09-608-285A-42
9	170.8	8.1	14747	4	US-09-557-800C-42
10	170.8	8.1	15977	4	US-09-608-285A-59
11	170.2	8.0	44848	4	US-09-435-739-42
12	170	8.0	685	4	US-09-227-357-100
13	169.8	8.0	13205	4	US-09-835-811-3
14	169.8	8.0	98844	4	US-09-791-211-10
15	169.6	8.0	63000	4	US-09-608-172-18
16	169.6	8.0	75395	4	US-09-984-890-3
17	167.8	7.9	63588	4	US-09-873-404-3
18	166.8	7.8	17000	4	US-09-679-299A-18
19	165.4	7.8	81001	4	US-09-750-580-1
20	164.8	7.8	6038	3	US-09-305-639-4
21	164.8	7.8	6038	4	US-09-525-160B-2
22	164.8	7.8	7622	3	US-09-305-639-1
23	164.8	7.8	7622	4	US-09-525-160B-1
24	164.4	7.8	436	4	US-09-621-976-19010
25	164.2	7.7	1519	4	US-09-716-129-45
26	163.4	7.7	45546	4	US-09-146-053-6
27	163.4	7.7	99500	4	US-09-798-096-10

28	163.2	7.7	3663	3	US-09-499-884-11	Sequence 11, Appl
29	163.2	7.7	43950	4	US-09-735-934A-3	Sequence 3, Appl
30	163.2	7.7	43950	4	US-10-060-332-3	Sequence 3, Appl
31	163	7.7	75395	4	US-09-984-890-3	Sequence 3, Appl
32	162.8	7.7	162450	4	US-09-345-882-1	Sequence 1, Appl
33	162.6	7.7	1332	4	US-09-584-568C-1	Sequence 1, Appl
34	162.2	7.6	12597	4	US-09-705-299-12	Sequence 12, Appl
35	162.2	7.6	29629	4	US-09-729-995-3	Sequence 3, Appl
36	162.2	7.6	29629	4	US-10-135-689-3	Sequence 3, Appl
37	162.2	7.6	35060	3	US-08-814-095-7	Sequence 7, Appl
38	161.6	7.6	90541	4	US-09-759-359A-3	Sequence 3, Appl
39	161.4	7.6	1426	4	US-09-620-312D-967	Sequence 967, App
40	161.4	7.6	99500	4	US-09-798-096-10	Sequence 10, Appl
41	161.2	7.6	6769	1	US-08-480-784-20	Sequence 20, Appl
42	161.2	7.6	6769	1	US-08-483-553-20	Sequence 20, Appl
43	161.2	7.6	6769	1	US-08-487-002-20	Sequence 20, Appl
44	161.2	7.6	6769	1	US-08-483-554B-20	Sequence 20, Appl
45	161.2	7.6	6769	1	US-08-488-011B-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1

US-09-724-864-32
; Sequence 32, Application US/09724864
; Patent No. 6380362
; GENERAL INFORMATION:
; APPLICANT: Watson, James D
; APPLICANT: Murison, James G.
; TITLE OF INVENTION: Polynucleotides, polypeptides expressed
; TITLE OF INVENTION: by the polynucleotides and methods for their use.
; FILE REFERENCE: 11000.105001
; CURRENT APPLICATION NUMBER: US/09/724,864
; PRIOR FILING DATE: 2000-11-28
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 1353
; TYPE: DNA
; ORGANISM: Mouse
US-09-724-864-32

Query Match 34.6%; Score 733; DB 4; Length 1353;
Best Local Similarity 76.9%; Pred. No. 3.9e-185;
Matches 1006; Conservative 0; Mismatches 275; Indels 28; Gaps 8;

QY	1	GAGCTCCCTCAGGAGCGGTTAGCTTCACTTCGCGAGGAGGCGGAGCTTCT	60
Db	46	GAGCGCTCCCGAGGAGCGGCGGTGAACCTTCCAGCAAGAGGCGTGTGCT	105
QY	61	CGCAGCGCGCAG-GGCGGCGCCAGGATCATGTCCACCACCATGCCAAGTGTGGCG	119
Db	106	CTTGAAGCTTCGCGCCAGCAGCTGAAGCATGGCCACCCAGCTGCAGGTGAGGG	165
QY	120	TTCTCTCTGTCATCTTGGGGTGGCGGCTGCATCGGCGCCACCGGATGACATGTGG	179
Db	166	CTTCTCTGTCCTCTCTGGGTCTGCGCGGCTGCATAGCGCCACTGGGATGACATGTGG	225
QY	180	AGCACCAGGAGCTTACAGCAACCCGTCACCTCCGTTTCCAGTACGAAGGGCTCTGG	239
Db	226	AGCACTCAAGACCTGTATGCAACCCAGTCACCGCGGTGTCCAGCATGAAGGGCTCTGG	285
QY	240	AGGAGCTCGTGAGCAGAGTTTCAGGCTTCACCGATGAGAGCCCTATTTCACATCTTG	299
Db	286	AGGAGTTGCGTGCAACAGAGCTCGGGGTTCACCGAGTCCCGGCCACTTTCACCATCTG	345
QY	300	GGACTTCCAGCCATCTGCAGGCGAGTGCAGCCCTGATGATCGTAGGATCGTCTCGGT	359
Db	346	GGCCTTCCAGCCATCTGCAGAGCTGACAGCCCTGATGATCGTGGCATTTGTTCTGGGG	405

360 GCATTGGCTCTGGTATCATCTTTGGCCCTGAATGCATCCGATTCGGCAGCATGAG 419
406 GTATCGGTATCTCTGTCCTTCCCTTCCCTTGAATGCATTCGATTCGGTATGATGAT 465
420 GACTCTCCAAAGCAACATGACACTGACCTTCGGGATCATGTTCTCATGTTCTCAGTCTT 479
466 GACTCTCCAAAGCAACATGACACTGACCTTCGGGATCATGTTCTCATGTTCTCAGTCTT 525
480 TGTCAATATGTCGAGTCTGTCGTTTTCGCAACATGCTGGTGAATCACTTCTGATGTC 539
526 TGTCAATATGTCGAGTCTGTCGTTTTCGCAACATGCTGGTGAATCACTTCTGATGTC 585
540 ACAGCTAACATGTACA-----CGGCATGGGTGGATGTCGAGACTGTTTCAGACC 590
586 ACAGCTAACATGTACAGCGGATGGCGGCATGGGTGGATGTCGAGACTGTTTCAGACC 645
591 AGGTACACATTTGGTGGGCTCTGTTTGGTGGGCTGGGTGGTGGTGGGCTCACACTAAT 650
646 AGGTACACATTTGGTGGGCTCTGTTTGGTGGGCTGGGTGGTGGTGGGCTCACACTAAT 705
651 GGGGGTGTGATGTCGATCGCTCGCTGGCGGCTGGCCAGCAAGAAACCAACTACAAA 710
706 GGGGGTGTGATGTCGATCGCTCGCTGGCGGCTGGCCAGCAAGCAAGCAACTTCAA 765
711 GCGGTTTCTTATCATGCTCAGGCTCAGGCTGTCGCTTCAAGCTGGAGGCTTCAAGGCC 770
766 GCTGTGCTTACCATGCTCTGGCCAAATGTTGCTTACAGGCTGGAGGCTTAAAGGCC 825
771 AGCACTGGCTTGGTCCACACAAACAAAGAGATATACGATGGAGTGGCGGCACA 830
826 AGCACTGGCTTGGTCCACACAAACAAAGAGATATACGATGGAGTGGCGGCACA 885
831 GAGGAGGAGTACATCTTATCTTCCAGCAGCTATGTCGATGCTTAAAGCTTCT 890
886 GAAGAGGAGTACATCTTATCTTCCAGCAGCTATGTCGATGCTTAAAGCTTCT 945
891 CAGCAGCGGCGGAAGAAATCTCC-----GGAGGCTCACCCAAACAAAGAGATC 942
946 CAACTGTGTGAGGAGGAACTCTTCCCAAGAGAGCTCACCC-CAAAGCAACGGAGT 1004
943 CCACTAGATTTCTTCTGCTTTTGCATCAGCTGGAAGTTAGAAAGCTTCGATTCA 1002
1005 CTACCTTGTTCCTTGTGATTTCACTGACATCTGAAAGTTGGTAAAGCTTGAATTTCA 1064
1003 TCTTTGGAGCGCAAAATGCTTAGCTCA-CTCTCTCTCTTAAATATCCACATAA 1061
1065 TCCATAGGAGGCTAGACAGTCTTGCCACATGTCCTGCTTAAATATCCACATCA 1124
1062 AACAGCTGAGTTA---TTTATGAATPAGGCTATAGCTCACATTTTCAATCTCTATT 1118
1125 AACAGCTGAGTTATCGTTTATGATGTTAGAGGCTTAACTACATCTTACCTTACCCCTCT 1184
1119 CTTTCTTTTAAATATACT---TTCTACTCTGATGAGAGATGTTGTTTAACTCTCTCTC 1176
1185 GCTTTTACCTAGACTTTCTTTTATCTGTTGATGAGGCTTAACTACATCTTACCTTAC 1244
1177 ACATTTTATGATTTTACAGACTCTCCCTCTCTCTCTTCTTCTTAAATATCCATGATGA 1236
1245 ATACTTTAATGGTTTAAAGAACTTTC---CTTCTCTGTTTAAATATCCATGATGA 1301
1237 TCTATTTCAGCTTATCCCAAGAAACTTTTGAAGGAAAGAGTAGA 1285
1302 TCGATTTTCCAGCTTGACCAAGGAAATTTTAAAGGAAAGGAAAGGAAAGGAAAGG 1350

RESULT 2
US-09-820-924-3
; Sequence 3, Application US/09820924
; Patent No. 655351
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen M. et al
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001213
; CURRENT APPLICATION NUMBER: US/09/820,924
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 39982
; TYPE: DNA
; ORGANISM: Human
US-09-820-924-3

Query Match 8.2%; Score 174.4; DB 4; Length 39982;
Best Local Similarity 77.6%; Pred. No. 1.3e-35;
Matches 239; Conservative 0; Mismatches 61; Indels 8; Gaps 2;

QY 1784 CTCACACTGTAATCTTAGCAGCTTTGGAGGCTGAGGAGGAGGATCACTTGGAGCCGAGA 1843
Db 31724 CTCATGCTGTAATCTTAGCAGCTTTGGAGGCAAGGAGATGATGATTTCTTGGAGCCAGG 31783
QY 1844 AGTTGAGACTAGCTGGGCAACATGAGAGAGCC-CTGTCTCTACAAATACAGAGAGAA 1902
Db 31784 AGTTTGAACACAGCTGGGCAACATGAGAGAGCCCTCTCTTACAGGAAATAC-----A 31836
QY 1903 AAAATCAGCCAGTCACTGTCATACACCTGTAGTCCAGGATTCGGGAGGCTGAGGTG 1962
Db 31837 AAAATAGCTGGTGTGTAGTGCACACTGAGTCCAGCTACCTGGGAGGCTGAGGTG 31896
QY 1963 GGAGGATCACTTGAAGCCAGGAGGAGTGGGGCTGCGAGTGAGGCTGATCACCAGTCA 2022
Db 31897 GGAGGATCACTTGAAGCCAGGAGGAGTGGGGCTGCGAGTGAGGCTGATCACCAGTCA 31956
QY 2023 CTCAGGAGGAGTCACTAGCAGGATCTGTCTTAAATAATAATAATAATAATAATAATA 2082
Db 31957 CTCAGGAGGAGTCACTAGCAGGATCTGTCTTAAATAATAATAATAATAATAATAATA 32016
QY 2083 CAGCAAGT 2090
Db 32017 AAGACATT 32024

RESULT 3
US-09-608-285A-8/c
; Sequence 8, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205

```
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(288)
; NAME/KEY: exon
; LOCATION: (1281)..(1580)
; NAME/KEY: exon
; LOCATION: (1820)..(1855)
; NAME/KEY: exon
; LOCATION: (2467)..(2555)
; NAME/KEY: exon
; LOCATION: (2863)..(2942)
; NAME/KEY: exon
; LOCATION: (3889)..(3950)
; NAME/KEY: exon
; LOCATION: (4894)..(4995)
; NAME/KEY: exon
; LOCATION: (5847)..(5987)
; NAME/KEY: exon
; LOCATION: (6966)..(7138)
; NAME/KEY: exon
; LOCATION: (8556)..(9365)
; NAME/KEY: misc_feature
; LOCATION: (3409)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc_feature
; LOCATION: (9214)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc_feature
; LOCATION: (9303)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc_feature
; LOCATION: (9311)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
US-09-608-285A-8

Query Match
Best Local Similarity 8.1%; Score 170.8; DB 4; Length 9365;
Matches 223; Conservative 26; Mismatches 53; Indels 6; Gaps 3;

QY 1783 GCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAGGATCACTTGCAGCCAG 1842
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7714 GCTCAYRCCTRTAATCYAGCACTTTGGGAGGCGYAGGCGGATCACCTGAGRTCAG 7655
QY 1843 AAGTTGGAGACTAGCCTGGGCAACATGGAGAGCCCTGTCTTACAAAATACAGAGAGAA 1902
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7654 RAGTTVRAGMYAGCTGSCCAACATGGTGAAACCCCGTCTCTACTATAAATACA---AA 7598
QY 1903 AAAATCAGCAGCTATGTTGGGATACACCTGTAGTCCCGAGATTCGGGAGGCTGAGGTG 1962
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7597 AAAATTAGCCRGCGCTGGTGGGCGCRYGCTGTAAATCCAGCTACTYGGGAGGCTGAGGCA 7538
QY 1963 GGAGGATCACTTGGAGGAGGCTGGGCTGAGTGCAGTCAGCCATCACACCACTGCA 2022
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7537 GRAGAATCGCTTGAACCCA-GGAGGYGRAGGTGCGATGAGCYGAGTATGCA 7479
QY 2023 CTCACAGCCAGTGA--CATAGCAGATCCTGTCTAAAAAATAAATAAATAAATAAATAA 2080
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7478 CTCACAGCCYRGMACAGACRAGACTCCGCTCTCAAAAAATAAATAAATAAATAAATAA 7419
QY 2081 CACAGCAA 2088
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7418 TAAATAAA 7411

RESULT 4
US-09-350-836B-8/c
; Sequence 8, Application US/09350836B
; Patent No. 6387645
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/35761
; CURRENT APPLICATION NUMBER: US/09/350,836B
; CURRENT FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(288)
; NAME/KEY: exon
; LOCATION: (1281)..(1580)
; NAME/KEY: exon
; LOCATION: (1820)..(1855)
; NAME/KEY: exon
; LOCATION: (2467)..(2555)
; NAME/KEY: exon
; LOCATION: (2863)..(2942)
; NAME/KEY: exon
; LOCATION: (3889)..(3950)
; NAME/KEY: exon
; LOCATION: (4894)..(4995)
; NAME/KEY: exon
; LOCATION: (5847)..(5987)
; NAME/KEY: exon
; LOCATION: (6966)..(7138)
; NAME/KEY: exon
; LOCATION: (8556)..(9365)
; NAME/KEY: misc_feature
; LOCATION: (3409)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
; NAME/KEY: misc_feature
; LOCATION: (9214)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
; NAME/KEY: misc_feature
; LOCATION: (9303)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
; NAME/KEY: misc_feature
; LOCATION: (9311)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
US-09-350-836B-8

Query Match
Best Local Similarity 8.1%; Score 170.8; DB 4; Length 9365;
Matches 223; Conservative 26; Mismatches 53; Indels 6; Gaps 3;

QY 1783 GCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAGGATCACTTGCAGCCAG 1842
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7714 GCTCAYRCCTRTAATCYAGCACTTTGGGAGGCGYAGGCGGATCACCTGAGRTCAG 7655
QY 1843 AAGTTGGAGACTAGCCTGGGCAACATGGAGAGCCCTGTCTTACAAAATACAGAGAGAA 1902
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7654 RAGTTVRAGMYAGCTGSCCAACATGGTGAAACCCCGTCTCTACTATAAATACA---AA 7598
QY 1903 AAAATCAGCAGCTATGTTGGGATACACCTGTAGTCCCGAGATTCGGGAGGCTGAGGTG 1962
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7597 AAAATTAGCCRGCGCTGGTGGGCGCRYGCTGTAAATCCAGCTACTYGGGAGGCTGAGGCA 7538
QY 1963 GGAGGATCACTTGGAGGAGGCTGGGCTGAGTGCAGTCAGCCATCACACCACTGCA 2022
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7537 GRAGAATCGCTTGAACCCA-GGAGGYGRAGGTGCGATGAGCYGAGTATGCA 7479
QY 2023 CTCACAGCCAGTGA--CATAGCAGATCCTGTCTAAAAAATAAATAAATAAATAAATAA 2080
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7478 CTCACAGCCYRGMACAGACRAGACTCCGCTCTCAAAAAATAAATAAATAAATAAATAA 7419
QY 2081 CACAGCAA 2088
Db |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
7418 TAAATAAA 7411

RESULT 5
US-09-370-265-8/c
; Sequence 8, Application US/09370265
; Patent No. 6447771
; GENERAL INFORMATION:
```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

[illegible]

APPLICANT: Young, George

```

1  TITLE OF INVENTION: Methods and Materials relating to CD59-Like
2  Polypeptides
3  FILE REFERENCE: 28110/36457
4  CURRENT APPLICATION NUMBER: US/09/557,800C
5  CURRENT FILING DATE: 2000-04-25
6  PRIOR APPLICATION NUMBER: 09/481,238
7  PRIOR FILING DATE: 2000-01-11
8  PRIOR APPLICATION NUMBER: 09/370,265
9  PRIOR FILING DATE: 1999-08-09
10 PRIOR APPLICATION NUMBER: PCT/US99/16180
11 PRIOR FILING DATE: 1999-07-16
12 PRIOR APPLICATION NUMBER: 09/350836
13 PRIOR FILING DATE: 1999-07-09
14 PRIOR APPLICATION NUMBER: 09/273447
15 PRIOR FILING DATE: 1999-03-19
16 PRIOR APPLICATION NUMBER: 09/122449
17 PRIOR FILING DATE: 1998-07-24
18 PRIOR APPLICATION NUMBER: 09/244444
19 PRIOR FILING DATE: 1999-02-04
20 PRIOR APPLICATION NUMBER: 09/118,205
21 PRIOR FILING DATE: 1998-07-16
22 NUMBER OF SEQ ID NOS: 56
23 SOFTWARE: PatentIn Ver. 2.0
24 SEQ ID NO 42
25 LENGTH: 14747
26 TYPE: DNA
27 ORGANISM: Homo sapiens
28 FEATURE:
29 NAME/KEY: misc feature
30 LOCATION: (13641)
31 OTHER INFORMATION: n = adenosine or guanine or thymidine
32 US-09-557-800C-42

```

QY	1783	GCTCACACCTGTAATCCTTAGCACTTTGGAGGCTGAGAGGAAGATCACTTGAGCCCGAG	1842
Db	11074	CTCAAYCCCTTAATTCYAGCACATTTGGGAGCYGAGYGGGTGGATCACTTGAGRTCAG	11015
QY	1843	AAGTTCCGAGACTAGCTCGGCAACATGAGAGAGCCCTGTCTCTACAAAAATACAGAGAGAA	1902
Db	11014	RAGTTYRAGAMVAGCTGCGCAACATGTGTGAACCCCGTCTCTACTTAAATAATACA	10958
QY	1903	AAAATCAGCCAGTCAATGCTGGCATACACTGTAGTCCCAAGCATTCCTCGGAGGCGCTGAGGTG	1962
Db	10957	AAAAATTAGCCRGGCGTGTGGGCGCYGCTGTAAATCCAGCTACTTGGGAGGCTGAGGCA	10898
QY	1963	GGAGGATCACTTGACCCACGAGGAGGTGGGGGCTGCAGTGAGCCATGATCAACACATGCA	2022
Db	10897	GRAGAAATCGTTTGAACCCA-GAGGGYGRAGGYTGCACTGAGCYGCGATCRYGCCAYTGCA	10839
QY	2023	CTCCAGCCAGGTGA--CATAGCGAGATCCTGTCTAAAAAAATAAAAAATAAATATCGAA	2080
Db	10838	CTCCAGCCYGGGMAACAAGAGCAGACTCCGTCTCAAAAAAATAAATAAATAATAAATAA	10779
QY	2081	CACAGCAA	2088
Db	10778	TAAATAAA	10771

RESULT 10
US-09-608-285A-59/c
; Sequence 59, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES

```

1 / FILE REFERENCE: 28110/36570
2 / CURRENT APPLICATION NUMBER: US/09/608,285A
3 / CURRENT FILING DATE: 2000-06-30
4 / PRIOR APPLICATION NUMBER: 09/583,231
5 / PRIOR FILING DATE: 2000-05-26
6 / PRIOR APPLICATION NUMBER: 09/557,800
7 / PRIOR FILING DATE: 2000-04-25
8 / PRIOR APPLICATION NUMBER: 09/481,238
9 / PRIOR FILING DATE: 2000-01-11
10 / PRIOR APPLICATION NUMBER: 09/370,265
11 / PRIOR FILING DATE: 1999-08-09
12 / PRIOR APPLICATION NUMBER: PCT/US99/16180
13 / PRIOR FILING DATE: 1999-07-16
14 / PRIOR APPLICATION NUMBER: 09/350,836
15 / PRIOR FILING DATE: 1999-07-09
16 / PRIOR APPLICATION NUMBER: 09/273,447
17 / PRIOR FILING DATE: 1999-03-19
18 / PRIOR APPLICATION NUMBER: 09/244,444
19 / PRIOR FILING DATE: 1999-02-04
20 / PRIOR APPLICATION NUMBER: 09/122,449
21 / PRIOR FILING DATE: 1998-07-24
22 / PRIOR APPLICATION NUMBER: 09/118,205
23 / PRIOR FILING DATE: 1998-07-16
24 / NUMBER OF SEQ ID NOS: 60
25 / SOFTWARE: Patentin Ver. 2.0
26 / SEQ ID NO 59
27 / LENGTH: 15977
28 / TYPE: DNA
29 / ORGANISM: Homo sapiens
30 / FEATURE:
31 / OTHER INFORMATION: CD39-L4/L66 Gene Sequence
32 / NAME/KEY: CDS
33 / LOCATION: (245)..(461)
34 / NAME/KEY: CDS
35 / LOCATION: (1454)..(1533)
36 / NAME/KEY: CDS
37 / LOCATION: (2734)..(2877)
38 / NAME/KEY: CDS
39 / LOCATION: (4364)..(4439)
40 / NAME/KEY: CDS
41 / LOCATION: (4679)..(4714)
42 / NAME/KEY: CDS
43 / LOCATION: (5326)..(5414)
44 / NAME/KEY: CDS
45 / LOCATION: (5723)..(5802)
46 / NAME/KEY: CDS
47 / LOCATION: (6751)..(6812)
48 / NAME/KEY: CDS
49 / LOCATION: (7758)..(7859)
50 / NAME/KEY: CDS
51 / LOCATION: (8712)..(8852)
52 / NAME/KEY: CDS
53 / LOCATION: (9831)..(9887)
54 / NAME/KEY: CDS
55 / LOCATION: (11613)..(11728)
56 / NAME/KEY: CDS
57 / LOCATION: (13146)..(13691)
58 / NAME/KEY: CDS
59 / LOCATION: (15702)..(15839)
60 / NAME/KEY: misc.feature
61 / LOCATION: (14871)
62 / OTHER INFORMATION: n = a or c o r g o t
63 / US-09-608-285A-59

```

	Query Match	8.1%; Score 170.8; DB 4;	Length 15977;
	Best Local Similarity	72.4%;	Pred. No. 7.1e-35;
	Matches 223;	Conservative 26;	Mismatches 53; Indels 6; Gaps 3;
OY	1783	GTCACACCCTGTATCCTAGCACTTTGGGAGCGTGAGGAGAAGATCACCTTGAAGCCAG	1842
	: : :	: : :	: : :
Dd	12304	GCTCAVRCRTTAATVCYAGACTTTGGAGGCYGAGGYGGTGGTCACTTGAERTCAG	12245
OY	1843	AAGTTGAGACTAGCTTGGGCCAATGGAGAAGCCCTGTCTCTACAATAACAGAGAA	1902

[illegible]

```

RESULT 11
US-09-435-739-42/c
; Sequence 42, Application US/09435739
; Patent No. 6664105
; GENERAL INFORMATION:
; APPLICANT: Pecker, Iris
; APPLICANT: Vlodayvsky, Israel
; APPLICANT: Feinstein, Elena
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY
; TITLE OF INVENTION: EXPRESSION OF SAME IN GENETICALLY MODIFIED CELLS
; FILE REFERENCE: 00/20454
; CURRENT APPLICATION NUMBER: US/09/435, 739
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42
; LENGTH: 44848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-435-739-42

```

	Query Match	8.0%;	Score 170.2;	DB 4;	Length 44848;
	Best Local Similarity	75.4%;	Pred. No. 1.8e-34;		
	Matches 254;	Conservative 0;	Mismatches 73;	Indels 10;	Gaps 3;
QY	1783	GCTCACACTGTAATCTTAGCACTTTGGAGGCTGAGGAGGAGGATCACTTTAGCCCCAG	1842		
Db	23517	GCTCATGCTGTCTATCTTAGCACTTTGGAGGCGAGGCGCGATCACTTGAGCCCGAG			
QY	1843	AGATTTCGAGACTAGCCTGGGCAACATGAGAGAGCCCTCTCTTCAAAATTCAGAGAGAA	1902		
Db	23457	GAGTTCAAGACAGCCTGGGCAACATGCGAAACACATCTCTCAAAAAACA-----CA	23403		

1903	QY	AAATACAGCCAGTCATGTGTGGCATACACCTGTAGTCCAGCATTC- ---CGGAGGGCTGA	1958
	Db	AAATTTAGCCGGGCATGATGTTAAAGCACCTGTAGTCCAGCTAGCTACTTTGGGAGGCTGA	23433
1959	QY	GGTGGGAGGATCACTTTGAGCCACAGGAGTTTGGGGCTGCGAGTGGCCATGATCACACCAC	2018
23342	Db	GGTGGGAGATCACTGAGCCACAGGAGGTCAGGCTGCGAGTGGCCATGATCATGCGAT	23283
2019	QY	TGCATCTCCAGCC- AGGTGCATACATGCGAGATCCTCTCTAAAAAAATAAAAAATAAATATG	2077
23282	Db	TGCACAGCAGCCTGGGCAACAGAAATGAGACCTGTCTCAAAAAAATAAAAAACAA	23223
2078	QY	GAAACACAGCATGCTCTAGGAGATGAGTTAAACATAAT	2114
23222	Db	AAAAACAAAAAATAAAGTAAGAGAAAGAAAAATAATT	23186

RESULT 12
US-09-227-357-100
: Sequence 100, Application US/09227357

```

1 Patent No. 6342581
2 GENERAL INFORMATION:
3 APPLICANT: Fischer et al.
4 TITLE OF INVENTION: 123 Human Secreted Proteins
5 FILE REFERENCE: PZ0101
6 CURRENT APPLICATION NUMBER: US/09/227,357
7 CURRENT FILING DATE: 1999-01-08
8 EARLIER APPLICATION NUMBER: PCT/US98/13684
9 EARLIER FILING DATE: 1998-07-07
10 EARLIER APPLICATION NUMBER: 60/051,926
11 EARLIER FILING DATE: 1997-07-08
12 EARLIER APPLICATION NUMBER: 60/052,793
13 EARLIER FILING DATE: 1997-07-08
14 EARLIER APPLICATION NUMBER: 60/051,925
15 EARLIER FILING DATE: 1997-07-08
16 EARLIER APPLICATION NUMBER: 60/051,929
17 EARLIER FILING DATE: 1997-07-08
18 EARLIER APPLICATION NUMBER: 60/052,803
19 EARLIER FILING DATE: 1997-07-08
20 EARLIER APPLICATION NUMBER: 60/052,732
21 EARLIER FILING DATE: 1997-07-08
22 EARLIER APPLICATION NUMBER: 60/051,931
23 EARLIER FILING DATE: 1997-07-08
24 EARLIER APPLICATION NUMBER: 60/051,932
25 EARLIER FILING DATE: 1997-07-08
26 EARLIER APPLICATION NUMBER: 60/051,916
27 EARLIER FILING DATE: 1997-07-08
28 EARLIER APPLICATION NUMBER: 60/051,930
29 EARLIER FILING DATE: 1997-07-08
30 EARLIER APPLICATION NUMBER: 60/051,918
31 EARLIER FILING DATE: 1997-07-08
32 EARLIER APPLICATION NUMBER: 60/051,920
33 EARLIER FILING DATE: 1997-07-08
34 EARLIER APPLICATION NUMBER: 60/052,733
35 EARLIER FILING DATE: 1997-07-08
36 EARLIER APPLICATION NUMBER: 60/052,795
37 EARLIER FILING DATE: 1997-07-08
38 EARLIER APPLICATION NUMBER: 60/051,919
39 EARLIER FILING DATE: 1997-07-08
40 EARLIER APPLICATION NUMBER: 60/051,928
41 EARLIER FILING DATE: 1997-07-08
42 EARLIER APPLICATION NUMBER: 60/055,722
43 EARLIER FILING DATE: 1997-08-18
44 EARLIER APPLICATION NUMBER: 60/055,723
45 EARLIER FILING DATE: 1997-08-18
46 EARLIER APPLICATION NUMBER: 60/055,948
47 EARLIER FILING DATE: 1997-08-18
48 EARLIER APPLICATION NUMBER: 60/055,949
49 EARLIER FILING DATE: 1997-08-18
50 EARLIER APPLICATION NUMBER: 60/055,953
51 EARLIER FILING DATE: 1997-08-18
52 EARLIER APPLICATION NUMBER: 60/055,950
53 EARLIER FILING DATE: 1997-08-18
54 EARLIER APPLICATION NUMBER: 60/055,947
55 EARLIER FILING DATE: 1997-08-18
56 EARLIER APPLICATION NUMBER: 60/055,964
57 EARLIER FILING DATE: 1997-08-18
58 EARLIER APPLICATION NUMBER: 60/056,360
59 EARLIER FILING DATE: 1997-08-18
60 EARLIER APPLICATION NUMBER: 60/055,684
61 EARLIER FILING DATE: 1997-08-18
62 EARLIER APPLICATION NUMBER: 60/055,984
63 EARLIER FILING DATE: 1997-08-18
64 EARLIER APPLICATION NUMBER: 60/055,954
65 EARLIER FILING DATE: 1997-08-18
66 EARLIER APPLICATION NUMBER: 60/058,785
67 EARLIER FILING DATE: 1997-09-12
68 EARLIER APPLICATION NUMBER: 60/058,664
69 EARLIER FILING DATE: 1997-09-12
70 EARLIER APPLICATION NUMBER: 60/058,660
71 EARLIER FILING DATE: 1997-09-12
72 EARLIER APPLICATION NUMBER: 60/058,661
73 EARLIER FILING DATE: 1997-09-12

```

```

; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 100
; LENGTH: 685
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-227-357-100

Query Match      8.0%; Score 170; DB 4; Length 685;
Best Local Similarity 79.9%; Pred. No. 1.9e-35;
Matches 238; Conservative 0; Mismatches 55; Indels 5; Gaps 3;

QY 1783 GCTCACACCTGTAATCCCTAGCACTTTGGGAGGCTGAGGAGGAGATCACTTGAGCCCGAG 1842
Db 392 GCTCATGCTGTAATCCCGACACTTTGGGAGGCTGAGGTGGGTGGATTCCTTGAGCTCAG 451

QY 1843 AAGTTTCGAGACTAGCTGGGCAACATGAGAGGCCCTGTCTCTACAAAATACAGAGAGAA 1902
Db 452 GAGTTCACAGCCGGCTTAGGCAACATGGTGAACCCCTGTCTCTATGAAAATACAG---A 508

QY 1903 AAAATCAGCCAGTCATGTTGGGCATACACCTGTAGTCCAGCATTCGGGAGGCTGAGGTG 1962
Db 509 AACATTAGCTGGTGTGGTGGCCACACACCTGTAGTCCAGCTACTTGGGAGGCTGAGGTG 568

QY 1963 GGAGGATCACTTGAGCCCGCAGGAGGTGTGGGCTGCGAGTGAAGCCATGATCACACCTGCA 2022
Db 569 GGAGGATCCCTTGAG-CTTGGGAGGTTGAGGCTGCGAGTGAAGTATACCACTGCA 627

QY 2023 CTCACGCC-AGGTGACATAGCGAGATCCTCTCTAAAAAAATAAAAAATAAATATGCA 2079
Db 628 TTCTAGCCTGGGTGACAGAGTGAGATGCTGTCTCCAAAAAATAAAAAAATACTGCA 685

RESULT 13
US-09-835-811-3
; Sequence 3, Application US/09835811
; Patent No. 6482936
; GENERAL INFORMATION:
; APPLICANT: HU, Song et al
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: US$ THEREOF
; FILE REFERENCE: CL0012228
; CURRENT APPLICATION NUMBER: US/09/835,811
; CURRENT FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 13205
; TYPE: DNA
; ORGANISM: Human
US-09-835-811-3

Query Match      8.0%; Score 169.8; DB 4; Length 13205;
Best Local Similarity 81.3%; Pred. No. 1.2e-34;
Matches 235; Conservative 0; Mismatches 47; Indels 7; Gaps 3;

QY 1783 GCTCACACCTGTAATCCCTAGCACTTTGGGAGGCTGAGGAGGAGATCACTTGAGCCCGAG 1842
Db 9973 GCTCACTCATGTAATCCAGCATTTTGGGAGGCTGAGGTGGGTGGATTCGCTTGAGCCCGAG 10032

QY 1843 AAGTTCGAGACTAGCTGGGCAACATGAGAGGCCCTGTCTCTACAAAATACAGAGAGAA 1902
Db 10033 GAGGTCAAGACAGCGCTGGGCAACATGCGGAACCCCTATTCTACAAAATA-----CA 10087

QY 1903 AAAATCAGCCAGTCATGTTGGGCATACACCTGTAGTCCAGCATTCGGGAGGCTGAGGTG 1962
Db 10088 AAATATAGCCAGGATGGTGGCATGCTACTGTAGTCCAGCTACTCAGAGAGCTGAGGTG 10147

QY 1963 GGAGATCACTTGAGCCCGCAGGAGGTGGGCTGCAAGTGAAGCCATGATCACACCTGCA 2022
Db 10148 GGAGATCACTTGAGCCCA-GGAGGTGAGGATGAGTGAAGTGTGATCACACCTGCA 10206

```


QY 2023 CTCAGCC-AGTGCATACAGGAGATCCTGTCTAAAAATAAATAAATA 2070
Db 10207 CTCAGCCTGGCAACAGAGACAGACCATCTCAAAAAAGAAAAA 10255

RESULT 14

US-09-791-211-10
; Sequence 10, Application US/09791211
; Patent No. 6448080
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION
; FILE REFERENCE: RTS-0205
; CURRENT APPLICATION NUMBER: US/09/791,211
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO 10
; LENGTH: 98844
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 24962
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 64383
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 65468
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 65469
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 65470
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 87130
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 89049
; OTHER INFORMATION: unknown
; OTHER INFORMATION:
US-09-791-211-10

Query Match 8.0%; Score 169.8; DB 4; Length 98844;
Best Local Similarity 81.3%; Pred. No. 3.7e-34;
Matches 235; Conservative 0; Mismatches 47; Indels 7; Gaps 3;

QY 1784 CTCACACCTGTAACTCTAGCATTGGGAGGCTGAGGAGGAGGATCATTGAGCCCGA 1843
Db 93465 CTCACACCTAAATCCAGCACTTTGGGAGGAGGAGGATCATTGAGCCCGA 93524
QY 1844 AGTTCGAGACTAGCCTGGGCAACATGGAGAGCCCTGTCTTACAAAAATACAGAGAA 1903
Db 93525 AGTTCGAGACAGCCTGGGCAACATGGAGATCCCTGTCTTACAAAAATA 93579
QY 1904 AAATCAGCAGTCATGGTGGCATACACCTGTAGTCCAGCAATCCGGAGGCTGAGGTGG 1963
Db 93580 AAATAGCTGGGCATGGTGGTCATGCTGTAGTCCAGCTATTGGGAGGCTGAGGTGG 93639
QY 1964 GAGGATCACTTGGCCCGAGGAGGTTGGGCTGAGTGGAGCCATGATCACACCTGCAC 2023
Db 93640 GAGGATCACTTGGAGCCCA-GGATGTTGGGCTGAGTGGAGTGGAGTGGAGTGGAGTGG 93698
QY 2024 TCCAGCCAGG-TGACATAGCAGATCCTGTCTTAAAAATAAATAAATA 2071
Db 93699 TCCAGGCTGGATGACAGAGCAGGCGCCGTTTCAAAAAAATAAATAAATA 93747

RESULT 15

US-09-780-172-18/c
; Sequence 18, Application US/09780172
; Patent No. 6607916
; GENERAL INFORMATION:
; APPLICANT: Robert McKay
; APPLICANT: Susan M. Freier
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF CASEIN KINASE 2-ALPHA EXPRESSION
; FILE REFERENCE: RTS-0159
; CURRENT APPLICATION NUMBER: US/09/780,172
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 96
; SEQ ID NO 18
; LENGTH: 63000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-780-172-18

Query Match 8.0%; Score 169.6; DB 4; Length 63000;
Best Local Similarity 74.0%; Pred. No. 3.2e-34;
Matches 228; Conservative 0; Mismatches 79; Indels 1; Gaps 1;

QY 1783 GCTCACACCTGTAACTCTAGCATTGGGAGGCTGAGGAGGAGGATCATTGAGCCCGA 1842
Db 10159 GCTCACACCTGTAACTCTAGCATTGGGAGGCTGAGGAGGAGGATCATTGAGCCCGA 10100
QY 1843 AAGTTCGAGACTAGCCTGGGCAACATGGAGAGCCCTGTCTTACAAAAATACAGAGAA 1902
Db 10099 GTGTTAGACACAGCCTGGGAAACACAGTGAACCCCATCTCTACTAAATACAAAAA 10040
QY 1903 AAATCAGCAGTCATGGTGGCATACACCTGTAGTCCAGCAATCCGGAGGCTGAGGTG 1962
Db 10039 AAAAAAGCCAGGCGTGGCAGCGTGGCTGTAGTCCAGCTACTTTGGGAGGCTGAGGCA 9980
QY 1963 GGAGGATCACTTGGCCCGAGGAGGTTGGGCTGAGTGGAGCCATGATCACACCTGCA 2022
Db 9979 GGAGGATTCCTTGAACCC-GGAGGTTGGAGGTTGAGTGGAGGCTGCTGCTGCTGCA 9921
QY 2023 CTCAGCCAGGTGACATAGCAGATCCTGTCTTAAAAATAAATAAATAAATAAATAA 2082
Db 9920 CTCAGCCTGGCAACAGAGTGGATGCTATCTCCAAAAAATAAATAAATAAATAA 9861
QY 2083 CAGCAAGT 2090
Db 9860 AAAAAAAT 9853

Search completed: September 2, 2004, 16:00:45
Job time : 121 secs

